# Large-scale Correlation of Mass and Galaxies with the Lyα Forest Transmitted Flux

Patrick McDonald[1,2], Jordi Miralda-Escudé [2], and Renyue Cen[3]

## ABSTRACT

We present predictions of the correlation between the Lyα forest absorption in quasar spectra and the mass within $\sim 5\,h^{-1}\,\mathrm{Mpc}$ (comoving) of the line of sight, using fully hydrodynamic and hydro-PM numerical simulations of the cold dark matter model supported by present observations. The observed correlation based on galaxies and the Lyα forest can be directly compared to our theoretical results, assuming that galaxies are linearly biased on large scales. Specifically, we predict the average value of the mass fluctuation, $< \delta_m >$, conditioned to a fixed value of the Lyα forest transmitted flux $\delta_F$, after they have been smoothed over a $10\,h^{-1}\,\mathrm{Mpc}$ cube and line of sight interval, respectively. We find that $< \delta_m > /\sigma_m$ as a function of $\delta_F/\sigma_F$ has a slope of 0.6 at this smoothing scale, where $\sigma_m$ and $\sigma_F$ are the rms dispersions (this slope should decrease with the smoothing scale). We show that this value is largely insensitive to the cosmological model and other Lyα forest parameters. Comparison of our predictions to observations should provide a fundamental test of our ideas on the nature of the Lyα forest and the distribution of galaxies, and can yield a measurement of the bias factor of any type of galaxies that are observed in the vicinity of Lyα forest lines of sight.

*Subject headings:* cosmology: theory—intergalactic medium—large-scale structure of universe—quasars: absorption lines

[1]Department of Physics, Princeton University, Princeton, NJ 08544; pmcdonal@feynman.princeton.edu

[2]Department of Astronomy, The Ohio State University, Columbus, OH 43210; jordi@astronomy.ohio-state.edu

[3]Princeton University Observatory, Princeton University, Princeton, NJ 08544; cen@astro.princeton.edu

# 1. INTRODUCTION

The prevalent theory to explain the Ly$\alpha$ forest is that the absorption lines arise from density variations in a photoionized intergalactic medium that originate in the gravitational evolution of primordial fluctuations. Both semi-analytic models and numerical simulations (Bi 1993; Cen et al. 1994; Zhang et al. 1995, 1998; Hernquist et al. 1996; Miralda-Escudé et al. 1996; Bi & Davidsen 1997) have shown that the predicted Ly$\alpha$ spectra appear remarkably similar to the observations. The good agreement of the predicted and observed flux distribution and power spectrum of the Ly$\alpha$ forest (Rauch et al. 1997; Croft et al. 1999; McDonald et al. 2000), and the large transverse size of the absorption systems (Bechtold et al. 1994; Dinshaw et al. 1994, 1997; Petitjean et al. 1998; Monier, Turnshek, & Hazard 1999; Dolan et al. 2000; López, Hagen, & Reimers 2000) are the basic tests that have so far been done and have supported the theory. In addition, this Ly$\alpha$ forest theory is derived from the general Cold Dark Matter (hereafter, CDM) model with parameters that are well constrained from several other observations (e.g., Primack 2000).

Another important test of our ideas of the Ly$\alpha$ forest can be done by observing the correlation of the transmitted flux in a spectrum with galaxies. Some of these observations have already been done for small scales and high column density absorbers, which have shown the expected strong correlation between galaxies and gas halos (e.g., Bergeron & Boissé 1991; Steidel, Dickinson, & Persson 1994; Lanzetta et al. 1995; Chen et al. 2001). Penton, Stocke, & Shull (2001) have probed this correlation at low redshift, and have found that the weak absorption lines are often found in low-density regions of the galaxy distribution. Recently, Adelberger et al. (2001) have carried out the first analysis of this correlation on large scales (several comoving Mpc) and at high redshift (using galaxies detected with the Lyman break technique), using the transmitted flux as the quantity to correlate with the mean number of galaxies in a specified region.

Inspired by the observational results of Adelberger et al. (2001), this paper presents detailed theoretical predictions for statistical functions similar to the one introduced by those authors. Namely, we analyze the mean value of the mass given an observed value of the transmitted flux, and the mean value of the transmitted flux for a fixed value of the mass, after both quantities have been smoothed over a certain region. The main difference between the functions we analyze and the function shown by Adelberger et al. (2001) is that the observed objects are of course galaxies, and our simulations predict only the distribution of the mass. However, the predictions of the simulations can still be compared to the galaxy observations to determine the relation between galaxies and mass, which is especially straightforward if linear bias is a sufficient description of the distribution of galaxies relative to the mass on the large scales being probed. We examine the dependence of these statistical functions on the various parameters affecting the Ly$\alpha$ forest in §3.

## 2. METHOD

We use several HPM (hydro-particle-mesh; see Gnedin & Hui 1998) simulations to model the Ly$\alpha$ forest, after testing that the HPM approximation is sufficient for our purposes by comparing to a fully hydrodynamic simulation. The standard HPM simulations that will be used most often in this paper have box size $40\,h^{-1}\,\mathrm{Mpc}$, with $512^3$ particles. The cosmological model is CDM with cosmological constant, in a flat universe with present matter density $\Omega_m = 0.4$, power spectrum index $n = 0.95$, and amplitude given by $\sigma_8 = 0.75$. All the results in this paper will be shown at $z = 3$, the typical redshift at which the observations have so far been done. The model we use is consistent with the observed Ly$\alpha$ forest power spectrum: at the typical Ly$\alpha$ forest scale of $k = 0.008\,(\mathrm{km\,s^{-1}})^{-1}$, it has $\Delta_\rho^2(k) = 0.29$ at $z = 3$ [where $\Delta_\rho^2(k)$ is the contribution per unit $\ln k$ to the variance of the linear theory mass density fluctuations], while the observed value is $\sim 0.26$ (McDonald et al. 2000; Croft et al. 2002). The redshift interval corresponding to $40\,h^{-1}\,\mathrm{Mpc}$ in this model near $z = 3$ is $\Delta z = 0.068$. The density-temperature relation assumed for the gas is $T(\Delta) = T_{1.4}(\Delta/1.4)^{\gamma-1}$, where $\Delta$ is the gas density divided by the mean gas density. We use the parameters $T_{1.4} = 17000$ K and $\gamma - 1 = 0.3$ in our standard simulation (the reason we specify the value of the temperature $T_{1.4}$ at $\Delta = 1.4$ is because the error in the temperature was determined to be smallest at this density in McDonald et al. 2002). Several other similar simulations are used where we vary each one of the relevant parameters (see §3), and we vary also the box size and resolution to verify the numerical convergence of the results. Unless otherwise indicated, all the results we show for a given set of parameters are averages of four simulations with different random initial perturbations.

### 2.1. Definitions

We define the quantity $\delta_F \equiv 1 - F/\bar{F}$, where $F$ is the fraction of transmitted flux in the Ly$\alpha$ forest after the spectrum has been smoothed along the line of sight with a filter, and $\bar{F}$ is the mean transmitted flux. We use a Gaussian smoothing filter in this paper, which is $W(kR) = \exp[-(kR)^2/2]$ in Fourier space. We also define $\delta_m$ as the redshift-space mass density perturbation, again smoothed over a specified region. We choose to smooth the mass fluctuation over cubes of $10\,h^{-1}\,\mathrm{Mpc}$ (comoving) in redshift space, centered on the Ly$\alpha$ forest line of sight, which is approximately the smoothing that is used in Adelberger et al. (2001). This is convenient to do in observations of galaxies, since galaxies are usually searched for in a square field of view, and they can simply be divided into redshift bins. Note that the size of our cubes in velocity units, $1280\,\mathrm{km\,s^{-1}}$, is more fundamental than the size in $h^{-1}\,\mathrm{Mpc}$, because the temperature and observed flux power spectrum are both measured in $\mathrm{km\,s^{-1}}$.

In general, the information that can be recovered from observations is the full joint probability distribution $P(\delta_F, \delta_g)$, where $\delta_g$ is the fluctuation in the number of galaxies. In this paper we present results from HPM numerical simulations for $<\delta_m|\delta_F>$, which is the mean value of the

mass fluctuation subject to the condition of a fixed value in the transmitted flux fluctuation, and for the analogous quantity $< \delta_F | \delta_m >$. The observed galaxies may not follow exactly the same fluctuations as the mass, even when smoothed over a scale of $10\,h^{-1}$ Mpc. However, if galaxies are linearly biased at this large scale, then $\delta_g = b\,\delta_m$, where $\delta_g$ is the fluctuation in the galaxies and $b$ is the bias factor. If we now define the quantity $\tilde{\delta}_m = \delta_m / \sigma_m$, and $\tilde{\delta}_g = \delta_g / \sigma_g$, where $\sigma_m$ and $\sigma_g$ are the rms fluctuation of $\delta_m$ and $\delta_g$, then $\tilde{\delta}_m = \tilde{\delta}_g$ and the correlation with $\delta_F$ is precisely predicted. For convenience, we also define the quantity $\tilde{\delta}_F = \delta_F / \sigma_F$, and we will present results for $< \tilde{\delta}_m | \tilde{\delta}_F >$ and $< \tilde{\delta}_F | \tilde{\delta}_m >$.

## 2.2. Smoothing

To choose a value of the smoothing radius $R$ for the smoothing filter of the Ly$\alpha$ forest transmitted flux, we compute first the correlation coefficient $< \delta_F \delta_m > / \sigma_F \sigma_m$ in our standard simulation, and show it as a function of $R$ in Figure 1. The correlation is maximum at $R = 3.5\,h^{-1}$ Mpc. The dependence of the correlation on $R$ is not surprising: if $R$ is too small, the value of $\delta_F$ is altered by small-scale fluctuations that are not related to the value of $\delta_m$ and therefore act as noise, and if $R$ is too big, then the fluctuations affecting $\delta_m$ are erased by smoothing in the value of $\delta_F$.

A Gaussian filter with $R = 3.5\,h^{-1}$ Mpc will be used for the smoothing of the Ly$\alpha$ spectrum throughout the rest of this paper. However, before proceeding, we examine how much our results differ if we use instead a top-hat smoothing for the Ly$\alpha$ spectrum, or if we vary the size of the cube over which the mass is smoothed.

Throughout the paper, our results will be presented as a set of four figures showing the four functions $< \delta_m | \delta_F >$, $< \delta_F | \delta_m >$, $< \tilde{\delta}_m | \tilde{\delta}_F >$, and $< \tilde{\delta}_F | \tilde{\delta}_m >$. In all these figures, the function $< \delta_m | \delta_F >$ has been obtained by creating spectra along each row of cells in the simulation, smoothing the spectra along the line of sight, selecting all the pixels in the spectra where the value of $\delta_F$ is inside a given bin, and computing the average $\delta_m$ for these pixels, where $\delta_m$ is computed in redshift-space. We use bins of width $\Delta \delta_F = 0.05$. Similarly, $< \delta_F | \delta_m >$ is calculated by averaging the values of $\delta_F$ for pixels in bins with a given $\delta_m$ (again using $\Delta \delta_m = 0.05$ for the bin width).

In Figure 2(a-d), we compare the results for our standard Gaussian filter (black, solid line) to the results obtained if we simply average the Ly$\alpha$ forest flux across the $10\,h^{-1}$ Mpc extent of the cube with which the mass is smoothed, i.e., use a top-hat filter with $R = 5\,h^{-1}$ Mpc (red, dotted line). The difference is not large, although noticeable, for the $\delta_m$-$\delta_F$ relations, and is practically zero for the $\tilde{\delta}_m$-$\tilde{\delta}_F$ relations.

We also show in Figure 2 the effect of changing the size of the cube used to smooth the mass density field from $10\,h^{-1}$ Mpc to $12\,h^{-1}$ Mpc (green, long-dashed line) or $8\,h^{-1}$ Mpc (blue, short-dashed line). In both of these cases we also change the smoothing length for the spectrum by the same factor. As the size of the smoothing cube is reduced, the mass fluctuations increase and they are more strongly correlated with the Ly$\alpha$ forest, as expected. We see that for $< \delta_m | \delta_F >$ it is
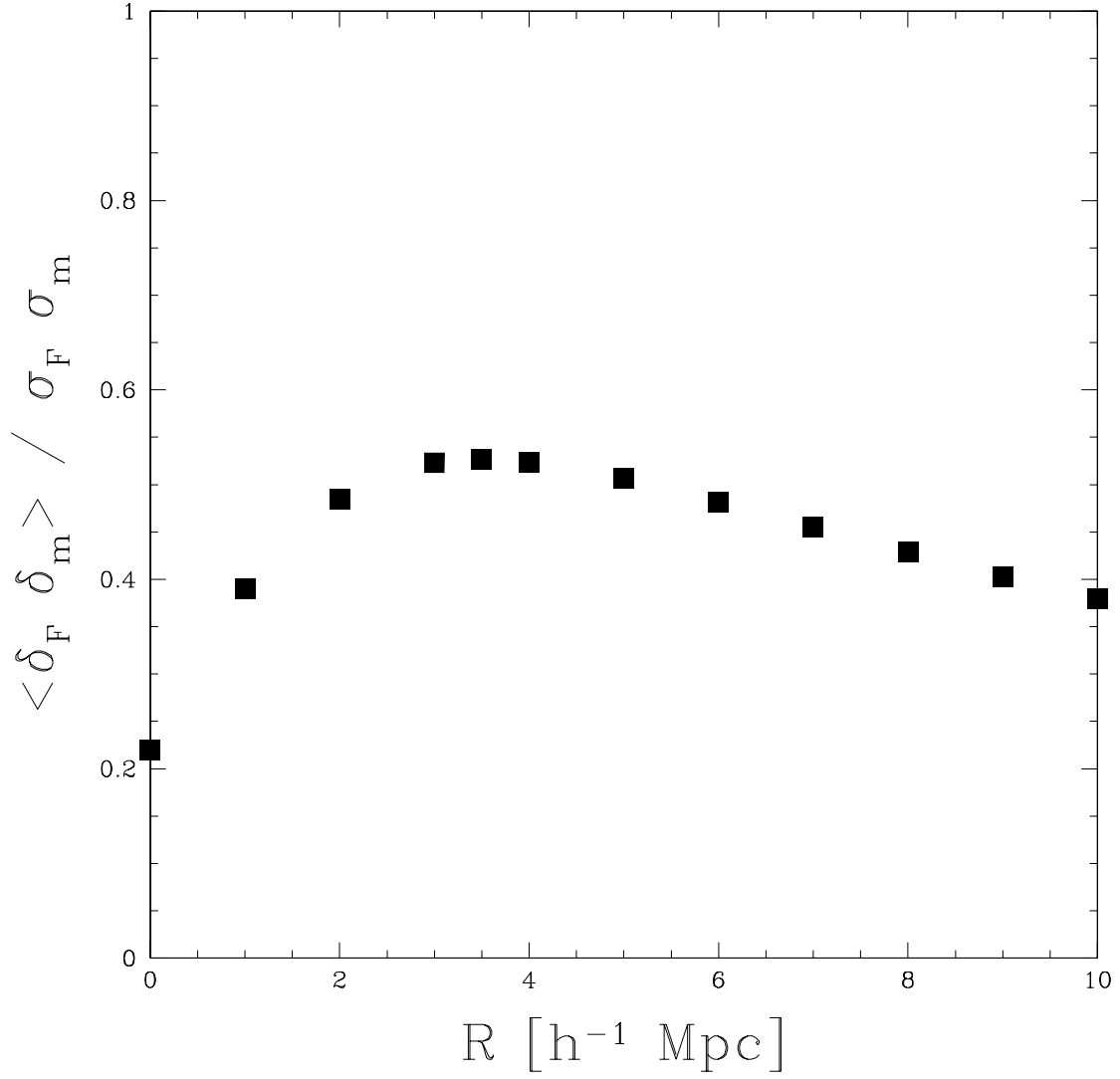
Fig. 1.— Correlation between mass and transmitted flux as a function of the smoothing applied on the spectrum, with the Gaussian filter, $W(kR) = \exp[-(kR)/2]$. The squares show the correlation, $<\delta_F \; \delta_m> /(\sigma_F \; \sigma_m)$.
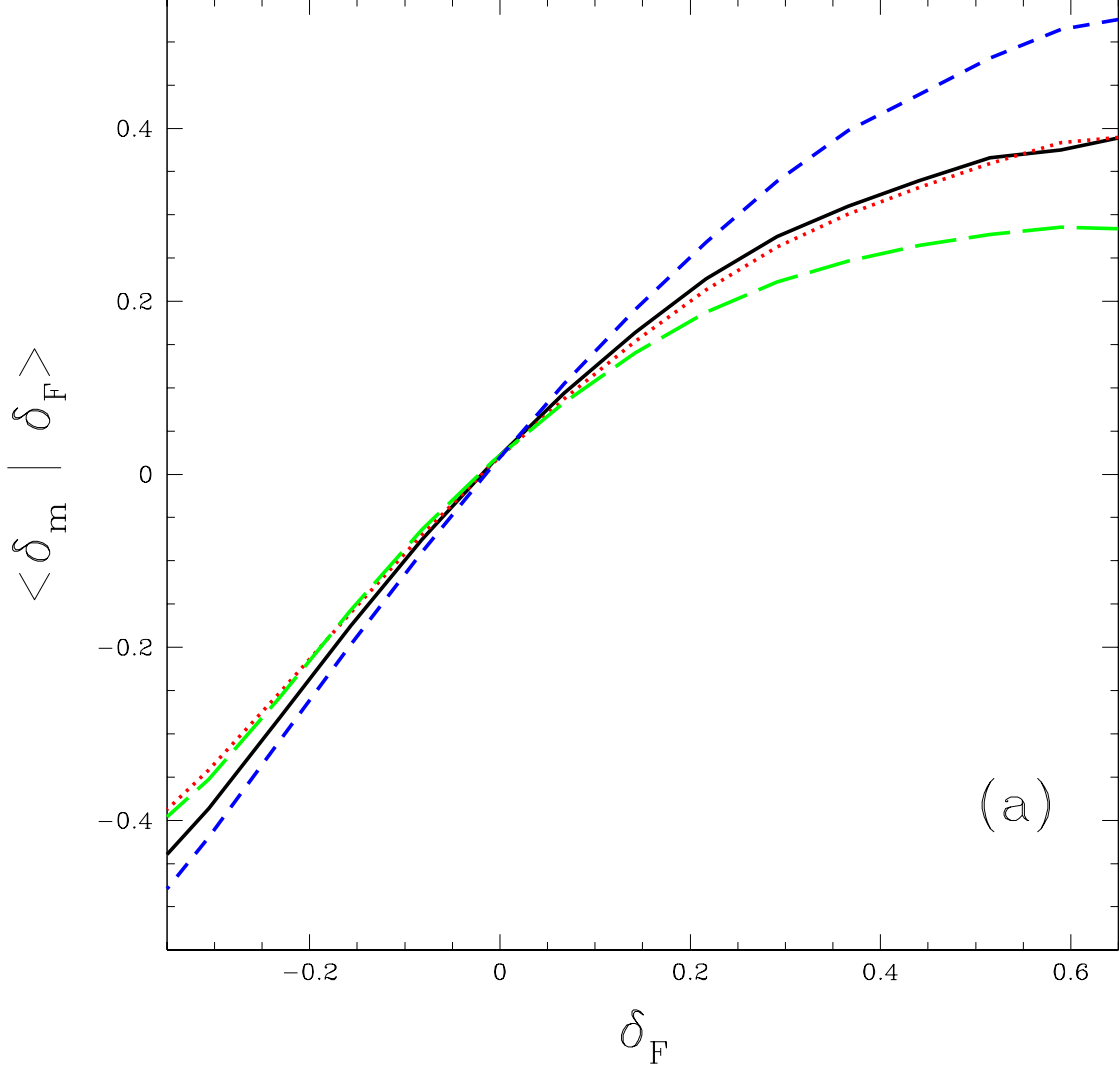
Fig. 2.— Comparison between our standard $3.5\,h^{-1}$ Mpc Gaussian smoothing of the spectrum with $10\,h^{-1}$ Mpc cube smoothing of the mass (*solid, black line*) and $5\,h^{-1}$ Mpc top-hat smoothing of the spectrum (*dotted, red line*), $12\,h^{-1}$ Mpc cube smoothing of the mass with $4.2\,h^{-1}$ Mpc Gaussian smoothing of the spectrum (*long-dashed, green line*), or $8\,h^{-1}$ Mpc cube smoothing of the mass with $2.7\,h^{-1}$ Mpc Gaussian smoothing of the spectrum (*short-dashed, blue line*). (a) mean mass as a function of transmitted flux, (b) mean flux as a function of mass, (c and d) are as (a and b) except that $\delta_F$ and $\delta_m$ are divided by their rms fluctuation. The mass is always computed in redshift space.

(b)
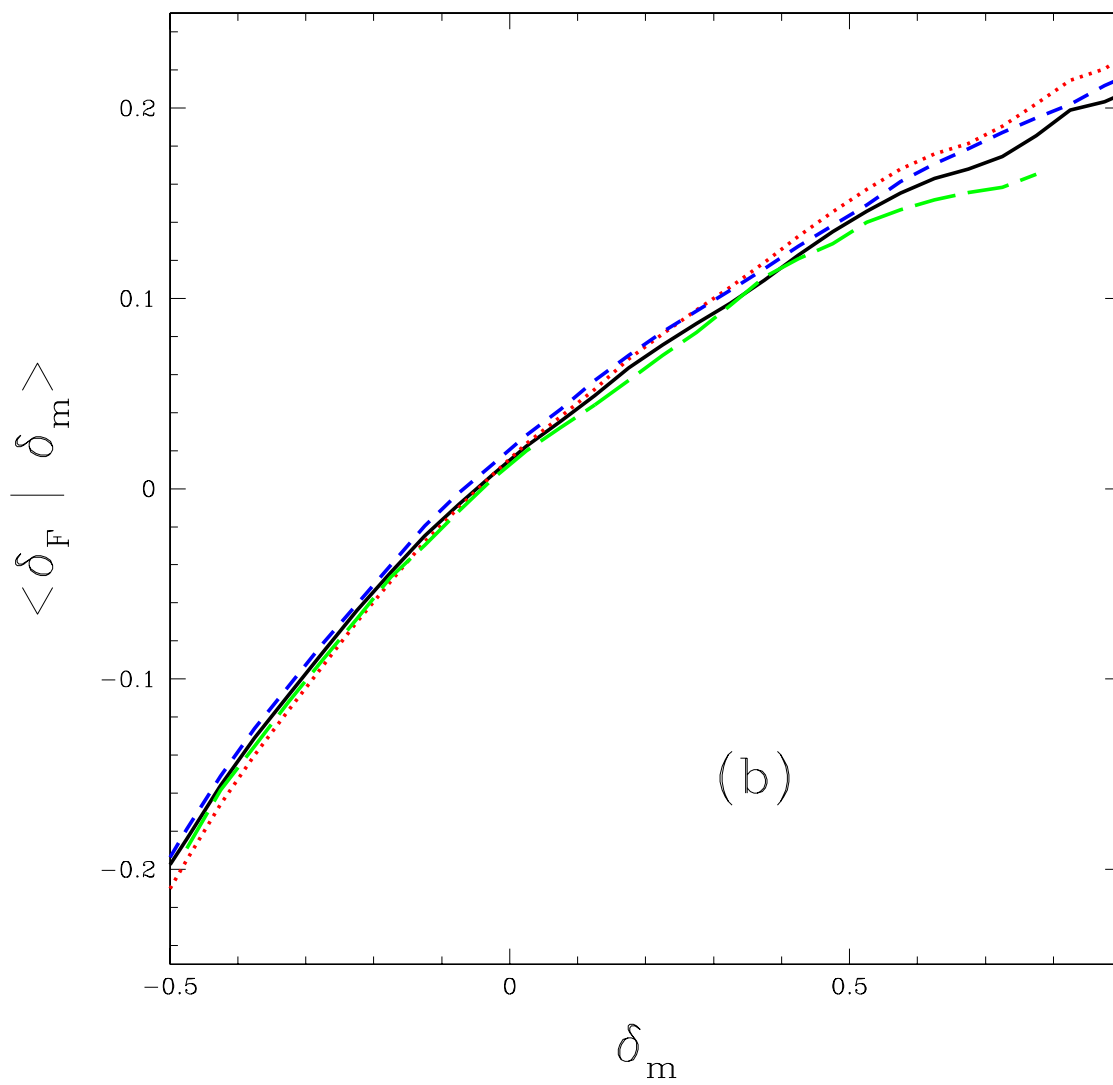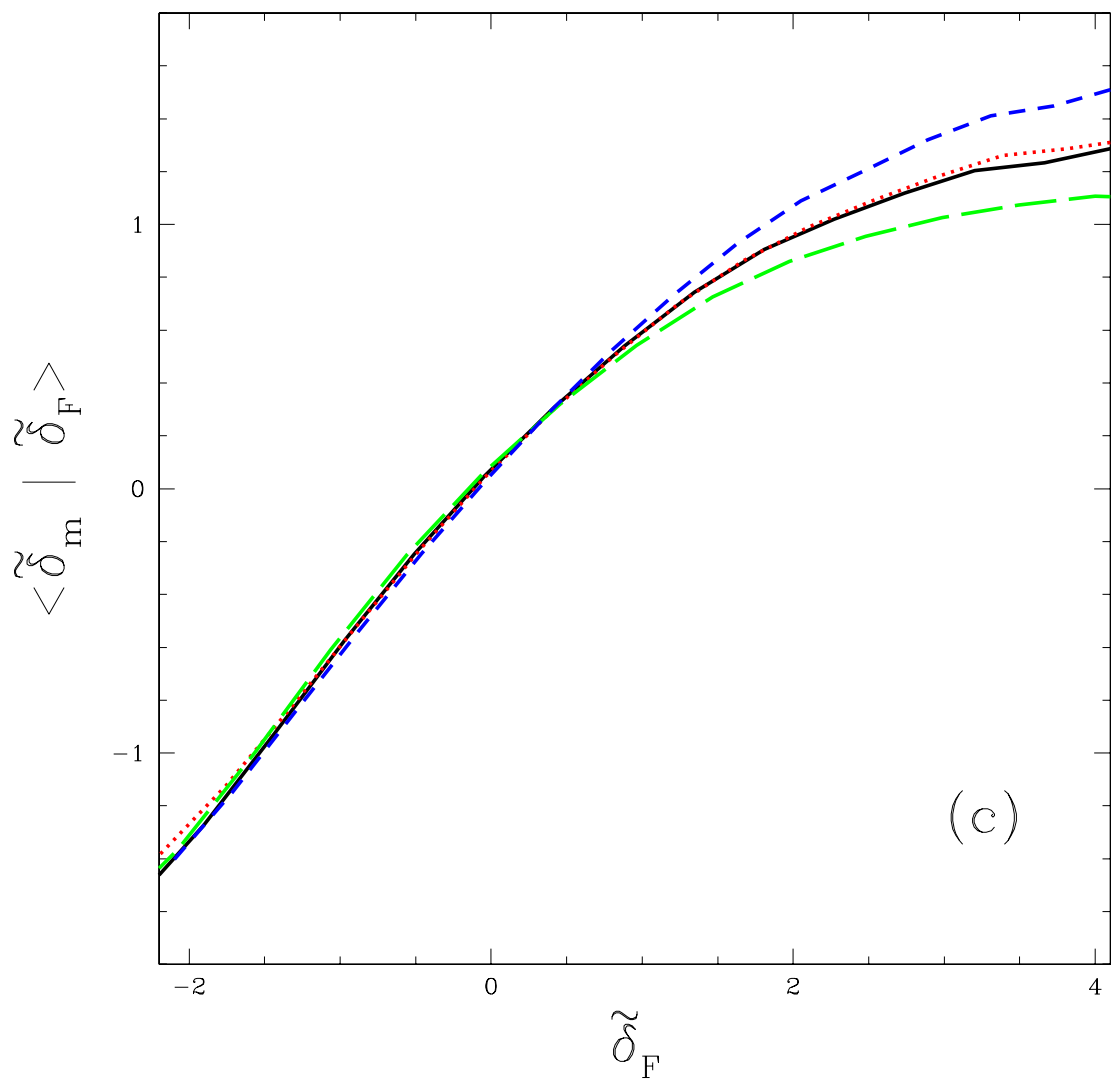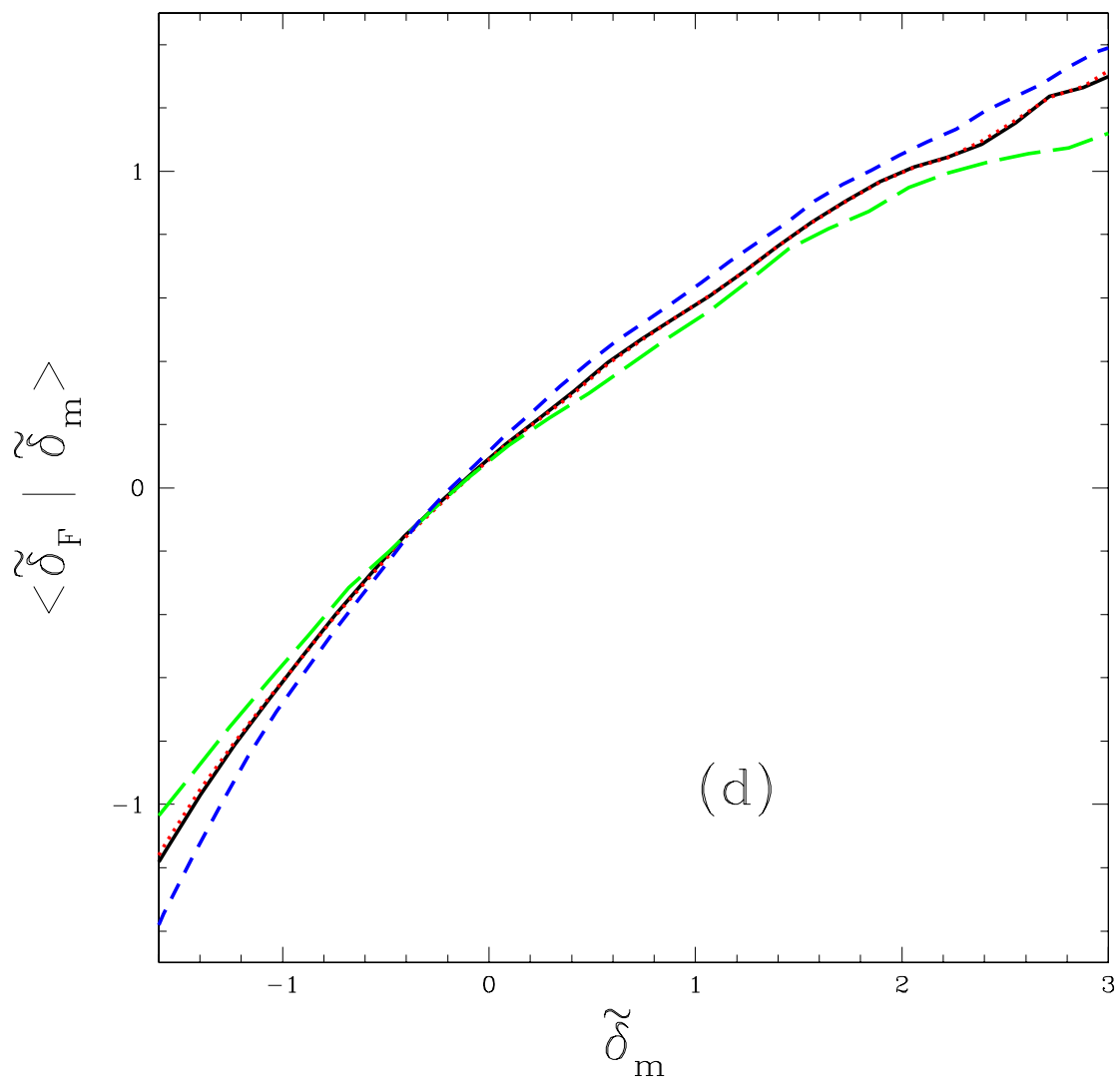
important to match the dimensions of the cube in the simulations and observations to compare the two. Note, however, that the transverse angular size of the cube depends on the cosmological geometry. The other plots show less sensitivity to the cube size.

To promote comparison between our results and observations or other simulations, in Table 1 we give $<\tilde{\delta}_m|\tilde{\delta}_F>$ and $<\tilde{\delta}_F|\tilde{\delta}_m>$ for the three cube sizes, along with $\sigma_F$ and $\sigma_m$ needed to convert from $\tilde{\delta}$ to $\delta$. The rows in the Table labeled 2s, 2ld, and 2sd correspond to cube size $10\,h^{-1}\,\mathrm{Mpc}$, $12\,h^{-1}\,\mathrm{Mpc}$, and $8\,h^{-1}\,\mathrm{Mpc}$, respectively.

## 2.3. Test of the HPM Approximation

We resort to approximate HPM simulations (Gnedin & Hui 1998) because it is impractical at the present time to perform fully hydrodynamic simulations of the required size for all of the parameter variations we would like to explore. However, we first test the accuracy of the HPM approximation by comparing to a state of the art hydrodynamic simulation of a similar cosmological model. This simulation is Eulerian, with box size $25\,h^{-1}\,\mathrm{Mpc}$ divided into $768^3$ cells for baryons, with $384^3$ dark matter particles (see Cen et al. 2001 for a more complete description). We compare to a $25\,h^{-1}\,\mathrm{Mpc}$, $512^3$ particle HPM simulation with identical initial Fourier modes up to the Nyquist frequency of the mesh (as we show in the next subsection, the resolution of these simulations is sufficient for convergence of our chosen statistics, so we do not need to worry about exactly how to equate resolution between the two).

The comparison for our statistics is shown in Figure 3, where the solid lines are the HPM results and the dotted lines are the fully hydrodynamic results. The agreement is excellent except in the highest $\delta_F$ regions in the $<\delta_m|\delta_F>$ and $<\tilde{\delta}_m|\tilde{\delta}_F>$ comparisons (a and c). For reference, in Figure 3(a and b) we plot the PDFs of $\delta_F$ and $\delta_m$, respectively (defined as the relative volume-weighted probability of finding $\delta$ within each of our usual bins). We see that the regions where the HPM approximation breaks down are extremely rare (at low $\delta_m$ in Figure 2b the lines terminate at the lowest density found in the simulations).

## 2.4. Convergence of the Results with Resolution and Box Size

We start by testing the sensitivity of our numerical results to the resolution. For this purpose, we use boxes with size $20\,h^{-1}\,\mathrm{Mpc}$ in addition to our standard value of $40\,h^{-1}\,\mathrm{Mpc}$.

The results for the four functions, $<\delta_m|\delta_F>$, $<\delta_F|\delta_m>$, $<\tilde{\delta}_m|\tilde{\delta}_F>$, and $<\tilde{\delta}_F|\tilde{\delta}_m>$, are shown in Figures 4(a,b,c,d) for our standard cosmological model, with box size $20\,h^{-1}\,\mathrm{Mpc}$, and resolution of $512^3$ and $256^3$ particles for the long-dashed and short-dashed lines, respectively. The solid and dotted lines show results with box size $40\,h^{-1}\,\mathrm{Mpc}$, again with $512^3$ and $256^3$ particles, respectively. The agreement between the $20\,h^{-1}\,\mathrm{Mpc}$ simulations differing only in the resolution
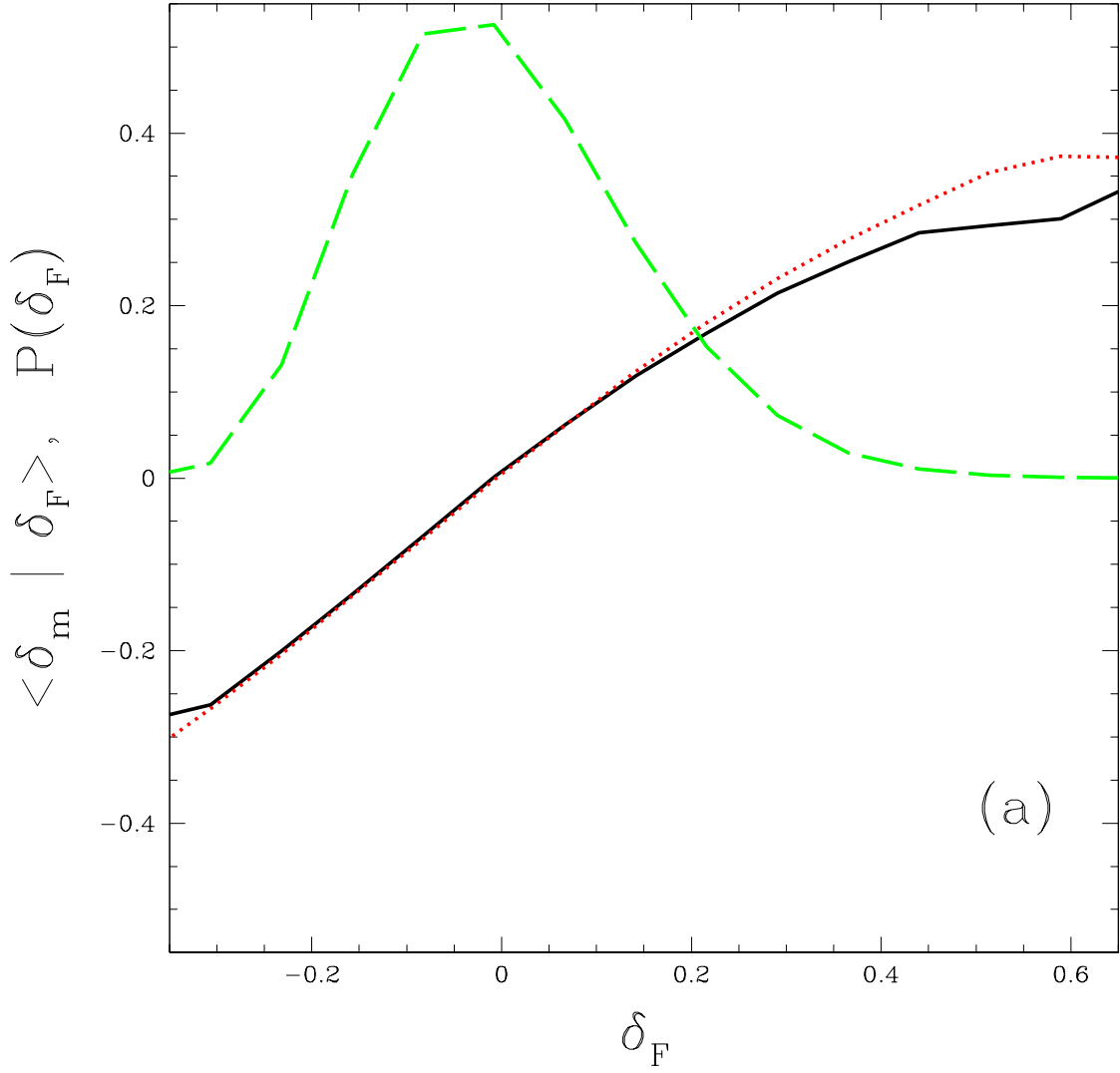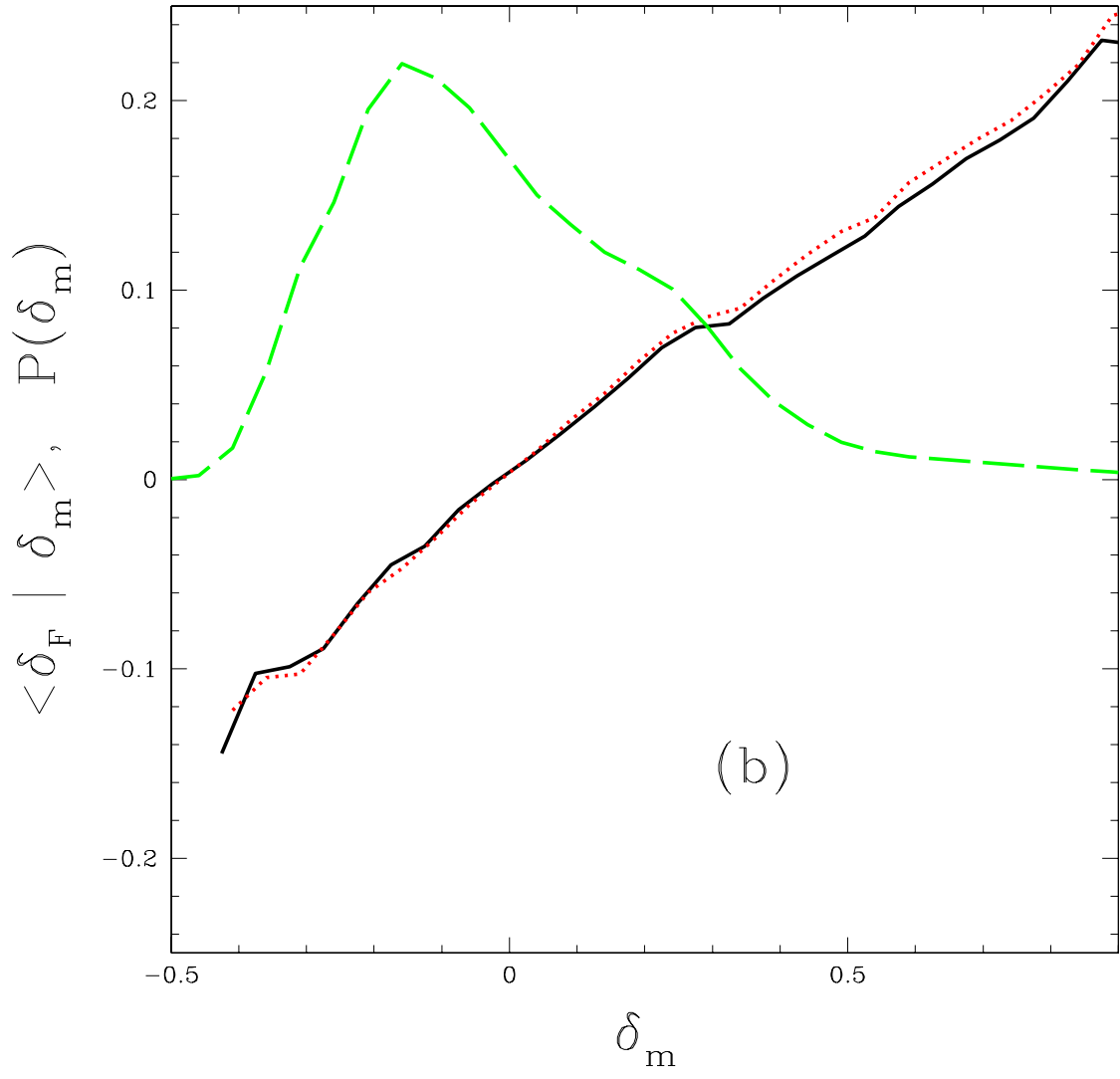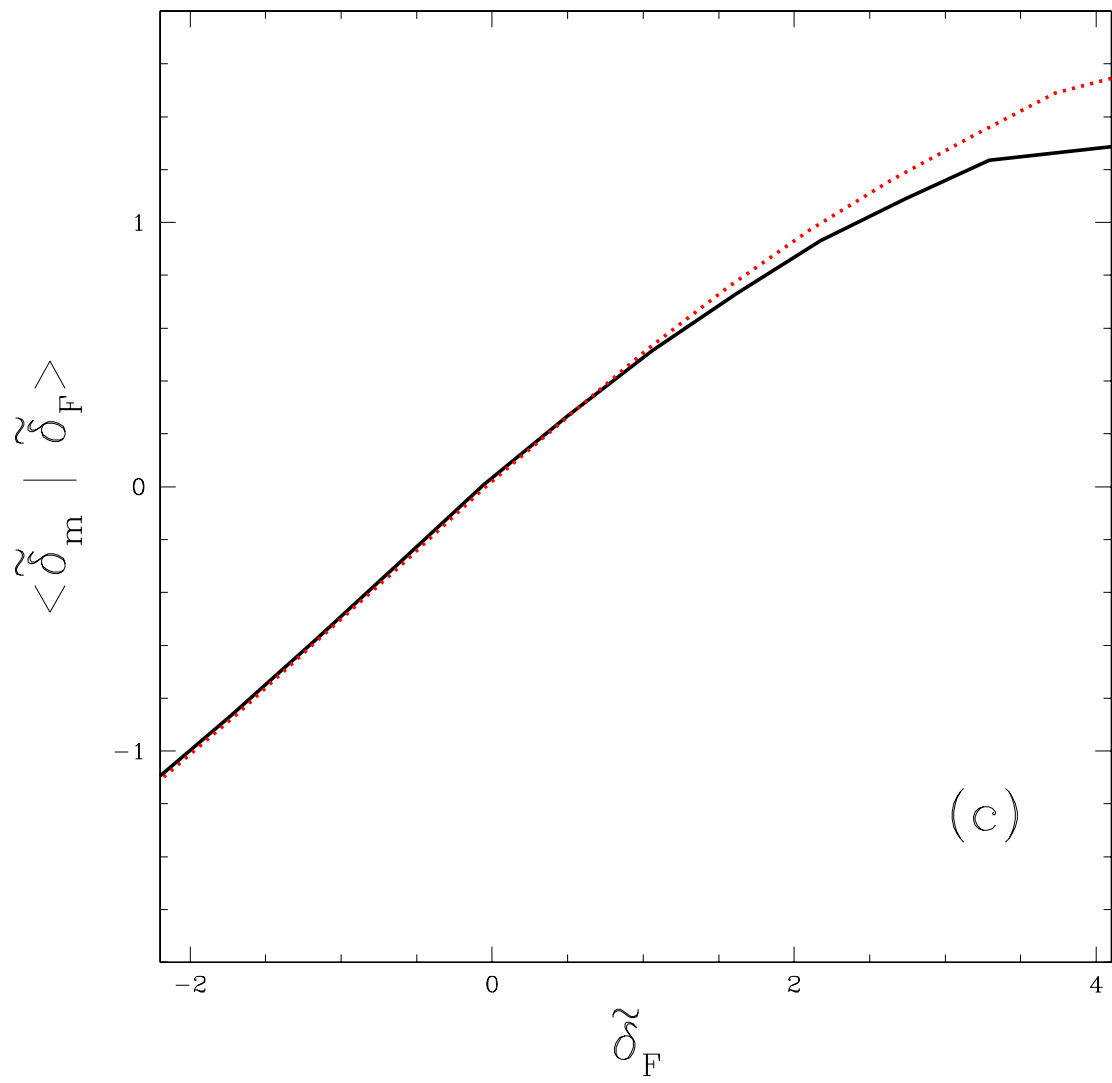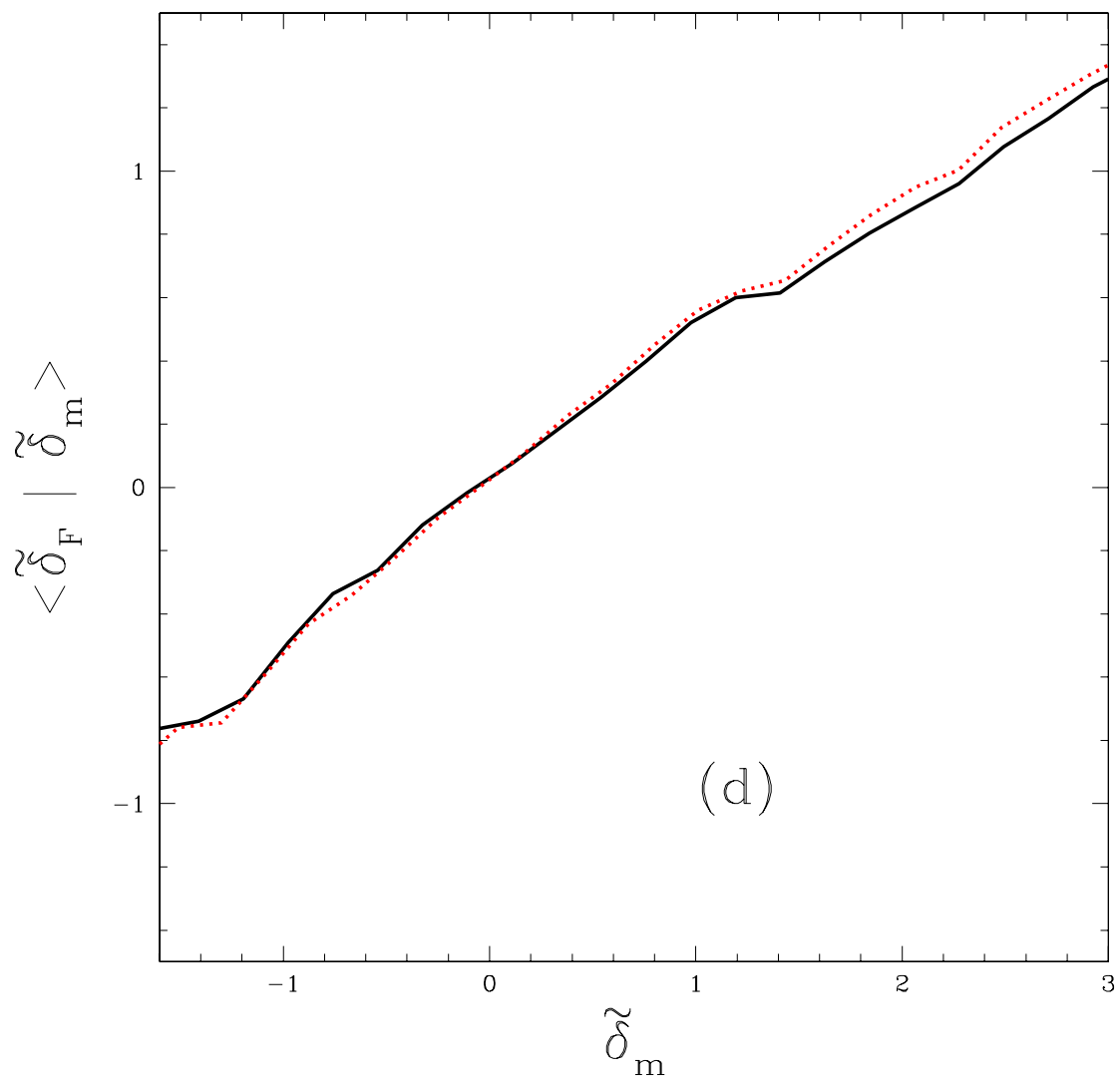
Fig. 3.— Test of the HPM approximation. *Black (solid) lines:* HPM results; *red (dotted) lines:* fully hydrodynamic results; *green (dashed) lines:* PDF of $\delta_F$ (a) and $\delta_m$ (b). The normalization of the PDFs is arbitrary.
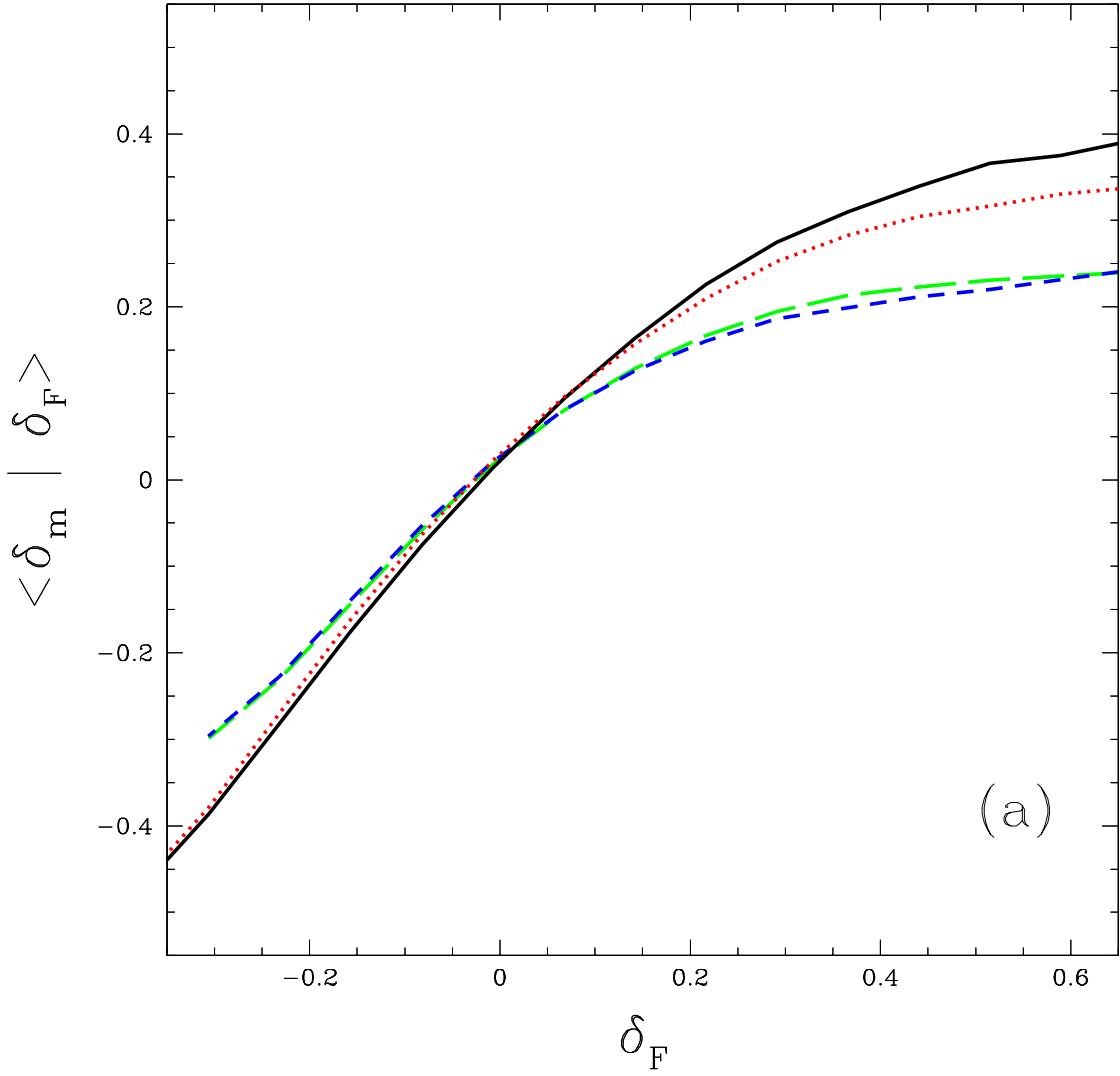
(b)

Fig. 4.— Resolution tests. *Green (long-dashed) line:* $20\,h^{-1}$ Mpc simulation box with $512^3$ particles; *blue (short-dashed) line:* $20\,h^{-1}$ Mpc, $256^3$ particles. *Black (solid) line:* $40\,h^{-1}$ Mpc, $512^3$ particles; *red (dotted) line:* $40\,h^{-1}$ Mpc, $256^3$ particles. (a) mean mass as a function of transmitted flux, (b) mean flux as a function of mass, (c and d) are as (a and b) except that $\delta_F$ and $\delta_m$ are divided by their rms fluctuation. The mass is always computed in redshift space.
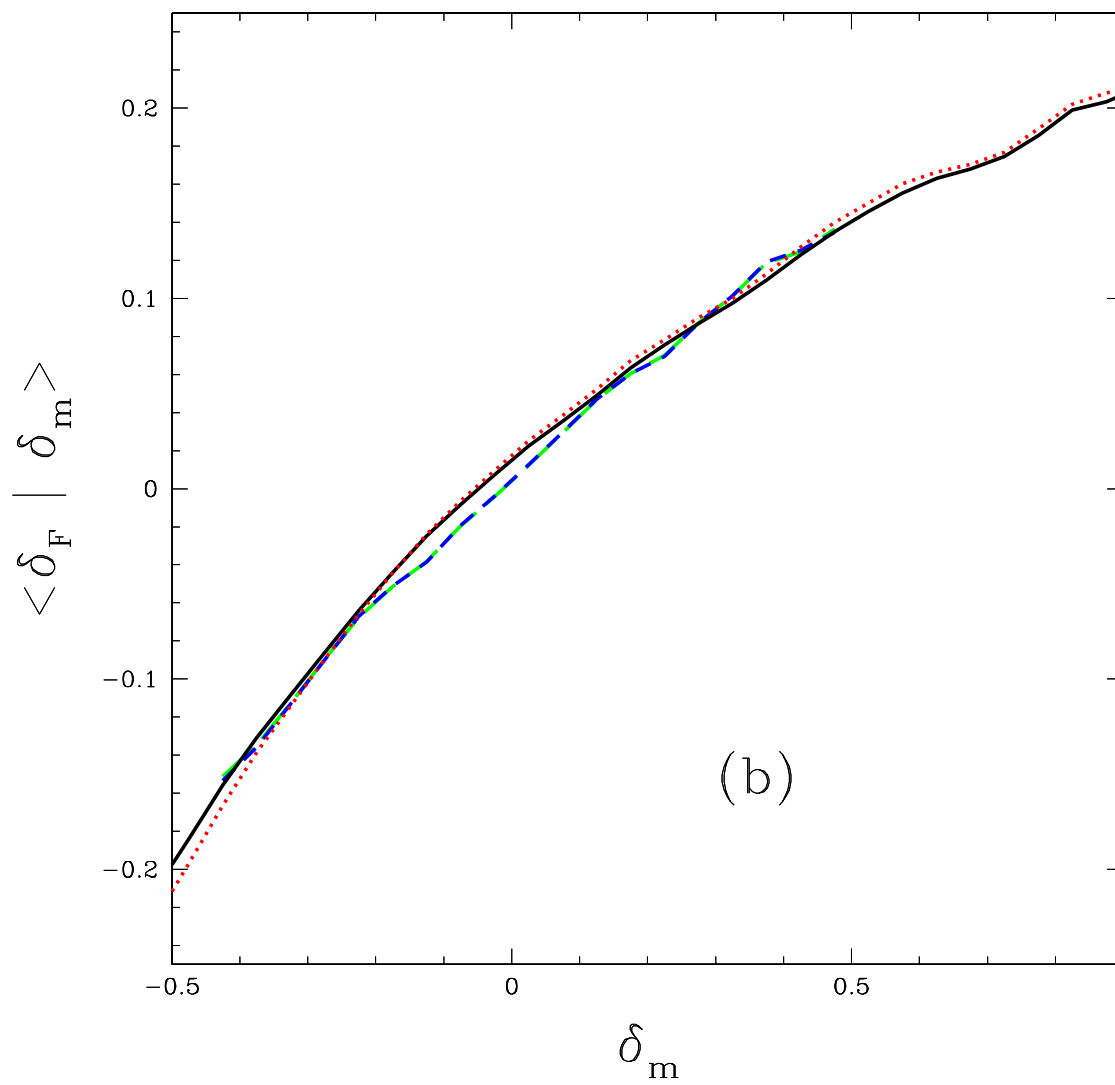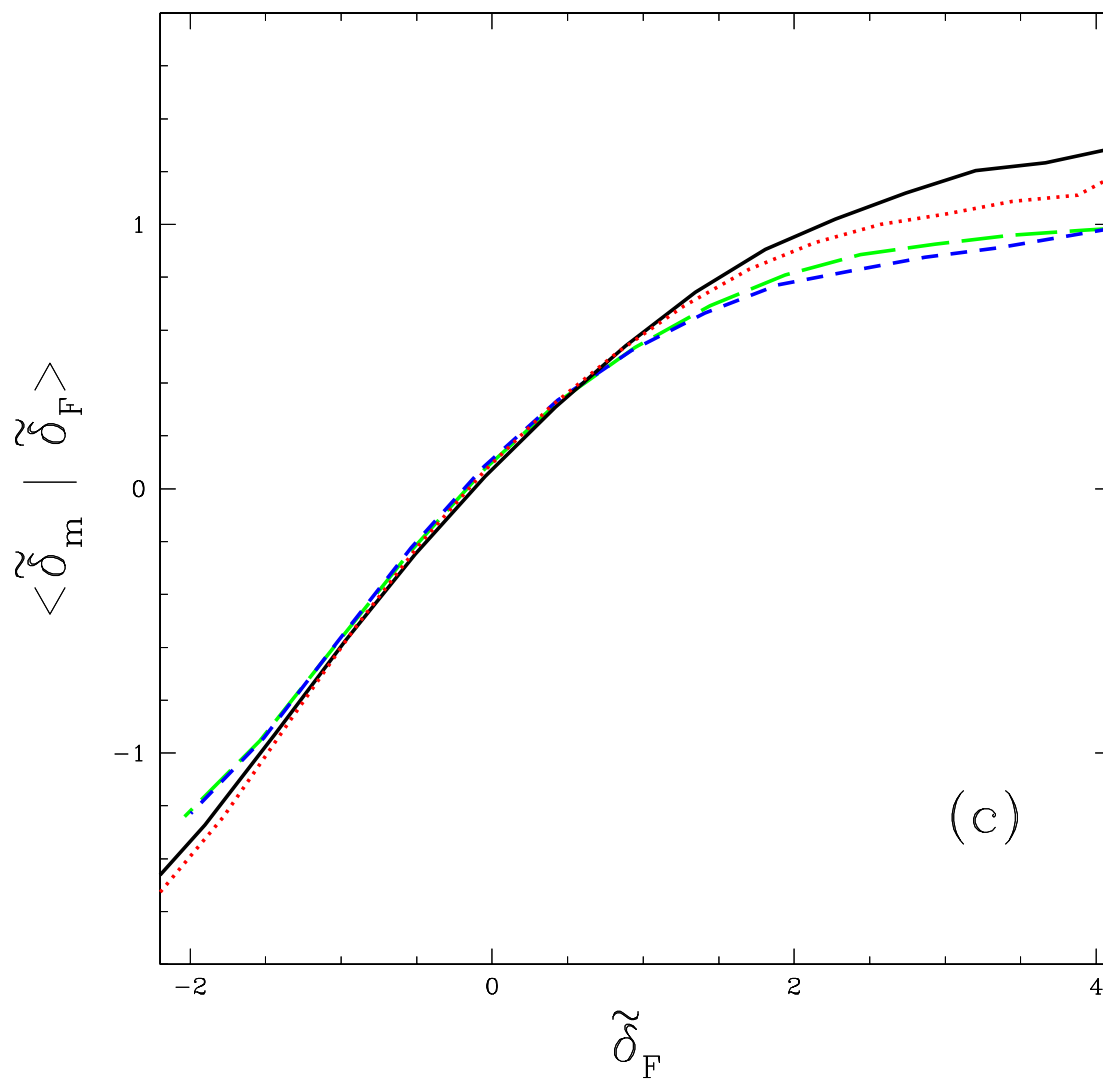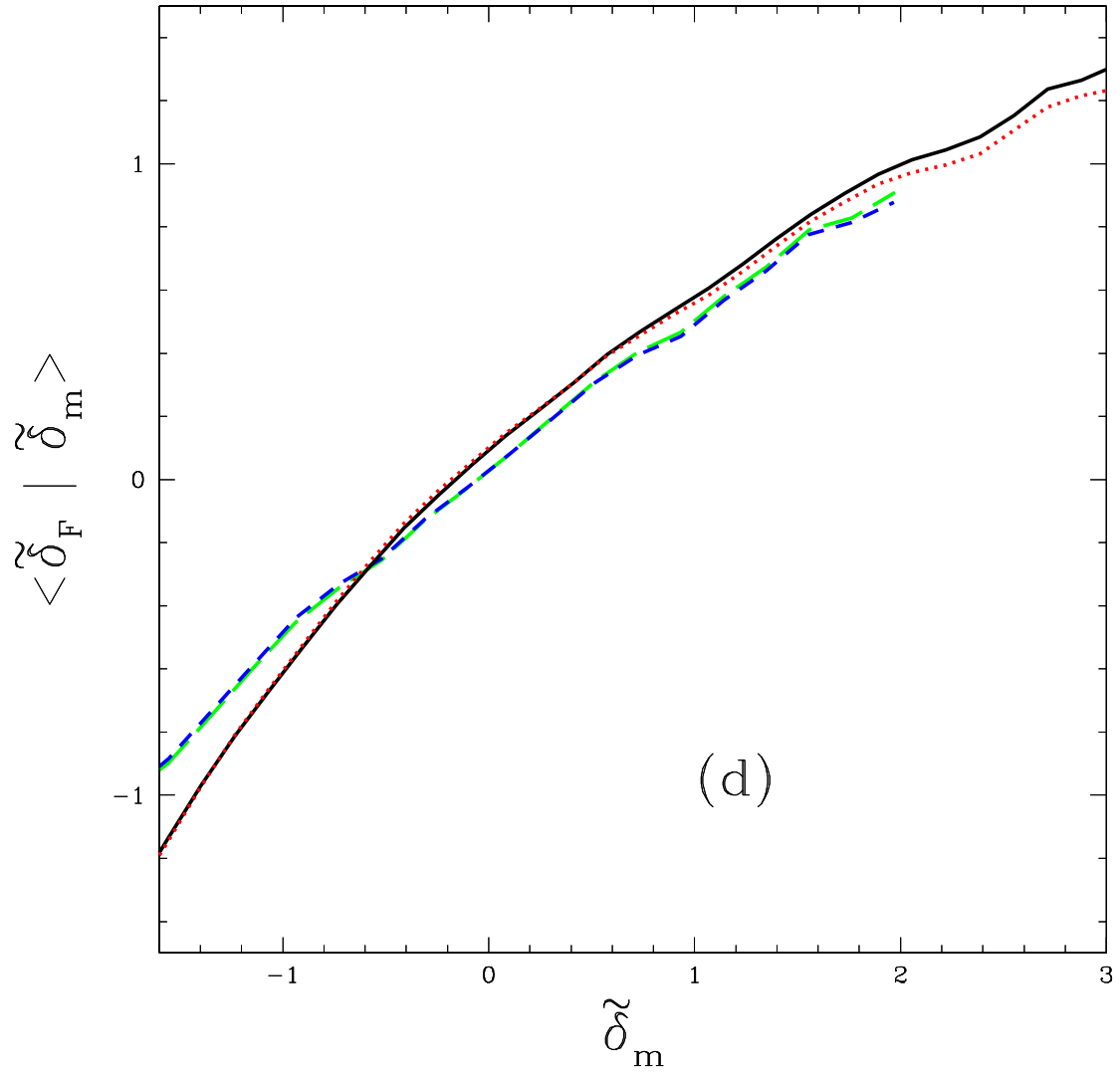
is excellent, verifying that $40\,h^{-1}\,\mathrm{Mpc}$ simulations with $512^3$ particles should be well resolved. The agreement between $40\,h^{-1}\,\mathrm{Mpc}$ simulations with different resolution is still quite good except at high $\delta_F$, indicating that $80\,h^{-1}\,\mathrm{Mpc}$, $512^3$ particle simulations could still be useful.

Figure 4 shows a disturbingly large difference between the results for $20\,h^{-1}\,\mathrm{Mpc}$ and $40\,h^{-1}\,\mathrm{Mpc}$ boxes; however, assessing the effect of the box size is more difficult than for the resolution, because we cannot use the same set of initial mode amplitudes at each wavenumber, so the statistical fluctuations in the simulations introduce significant random differences in the result. Figures 5(a-d) compare results for $40\,h^{-1}\,\mathrm{Mpc}$ boxes (*black, solid lines*), $20\,h^{-1}\,\mathrm{Mpc}$ boxes (*red, dotted lines*), and $80\,h^{-1}\,\mathrm{Mpc}$ boxes (*green, dashed lines*), all with $512^3$ particles. The thick lines show the average of four simulations, and the thin lines show the results from each separate simulation to demonstrate the statistical error in the curves.

Figure 5(a) shows good agreement between the 40 and $80\,h^{-1}\,\mathrm{Mpc}$ boxes for $<\delta_m|\delta_F>$, except at large $\delta_F$, where the trend toward increasing mean mass with increasing box size from 20 to 40 to $80\,h^{-1}\,\mathrm{Mpc}$ is probably not a result of statistical fluctuations. In fact, the true difference between the 40 and $80\,h^{-1}\,\mathrm{Mpc}$ boxes is even larger than what is shown because the decreased resolution in the $80\,h^{-1}\,\mathrm{Mpc}$ box has the effect of suppressing $<\delta_m|\delta_F>$ at large $\delta_F$ (see Figure 4a). It is clear that much of the flattening of $<\delta_m|\delta_F>$ that we see at large $\delta_F$ is an effect of the finite box size. Fortunately, the other statistics shown in Figure 5(b-d) generally exhibit better convergence than $<\delta_m|\delta_F>$.

The results in Figures 5(a-d) should also serve as a cautionary reminder that statistical fluctuations from one simulation to another can be large, and must be taken into account when comparing to observations or to other simulations, especially single, relatively small simulations. Surprisingly, the scatter between the separate simulations in the high $\delta_F$ corner of the plot is not reduced by increasing the box size from 40 to $80\,h^{-1}\,\mathrm{Mpc}$. This may be an indication that the statistical fluctuations in the result are still dominated by the longest wavelength modes in the box.

For reference, for our standard model, in a $40\,h^{-1}\,\mathrm{Mpc}$ box with our standard smoothing, the rms fluctuations in $\delta_F$ and $\delta_m$ are $\sigma_F = 0.16$ and $\sigma_m = 0.30$.

## 3. PARAMETER DEPENDENCE OF THE LY$\alpha$ FOREST - MASS CORRELATION

We now examine the dependence of the four functions $<\delta_m|\delta_F>$, $<\delta_F|\delta_m>$, $<\tilde{\delta}_m|\tilde{\delta}_F>$, and $<\tilde{\delta}_F|\tilde{\delta}_m>$ on the most important parameters determining the properties of the Ly$\alpha$ forest. These are the mean transmitted flux, the mean temperature-density relation of the intergalactic gas (parameterized as $T = T_0\Delta^{\gamma-1}$), and the amplitude and power-law slope of the power spectrum. These parameters fully determine all the physical effects that are incorporated in an HPM simulation and the calculation of the Ly$\alpha$ spectra: the underlying dark matter fluctuations (assumed to be Gaussian initially), the smoothing of the gas distribution on the Jeans scale, and the neutral
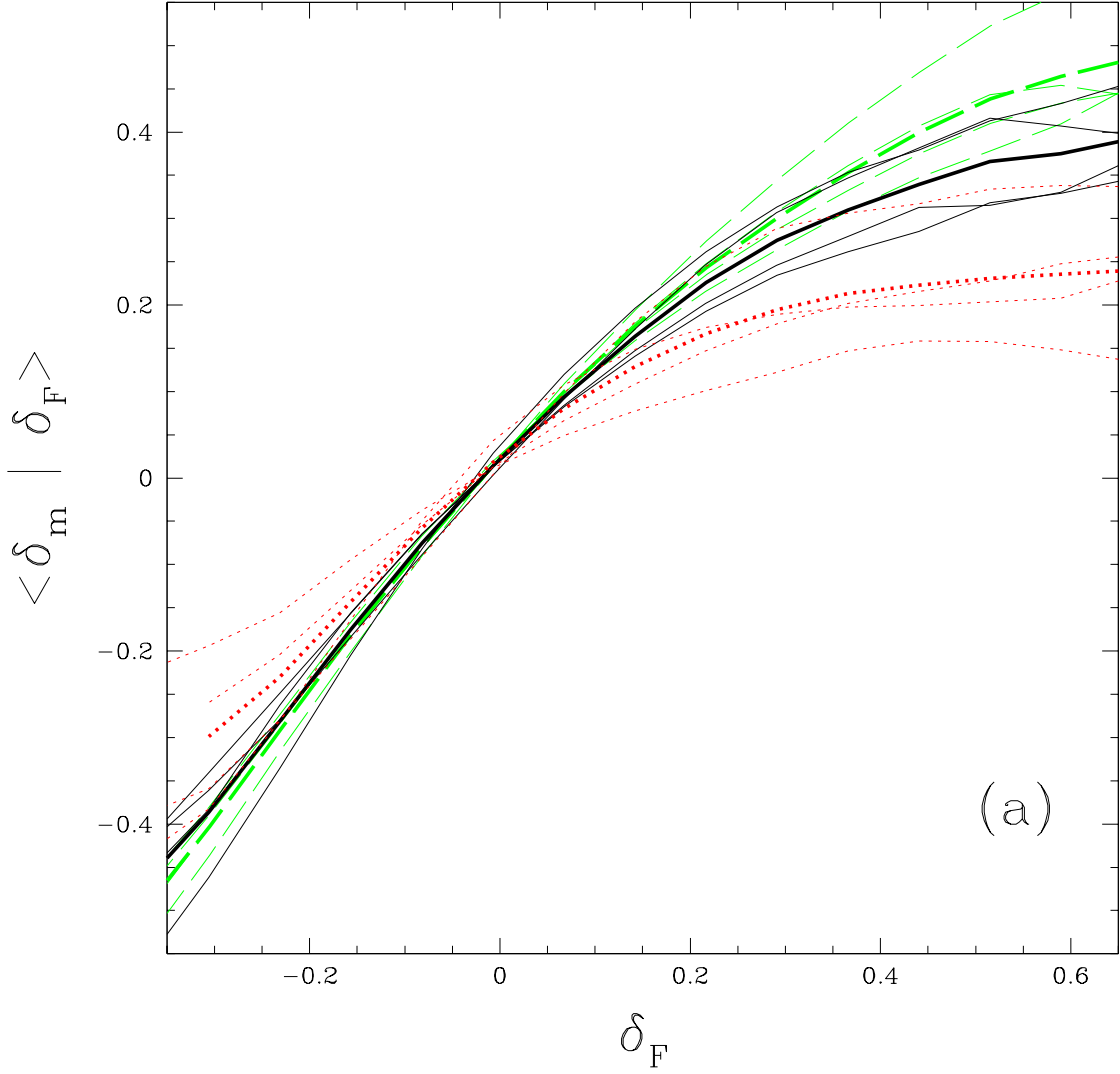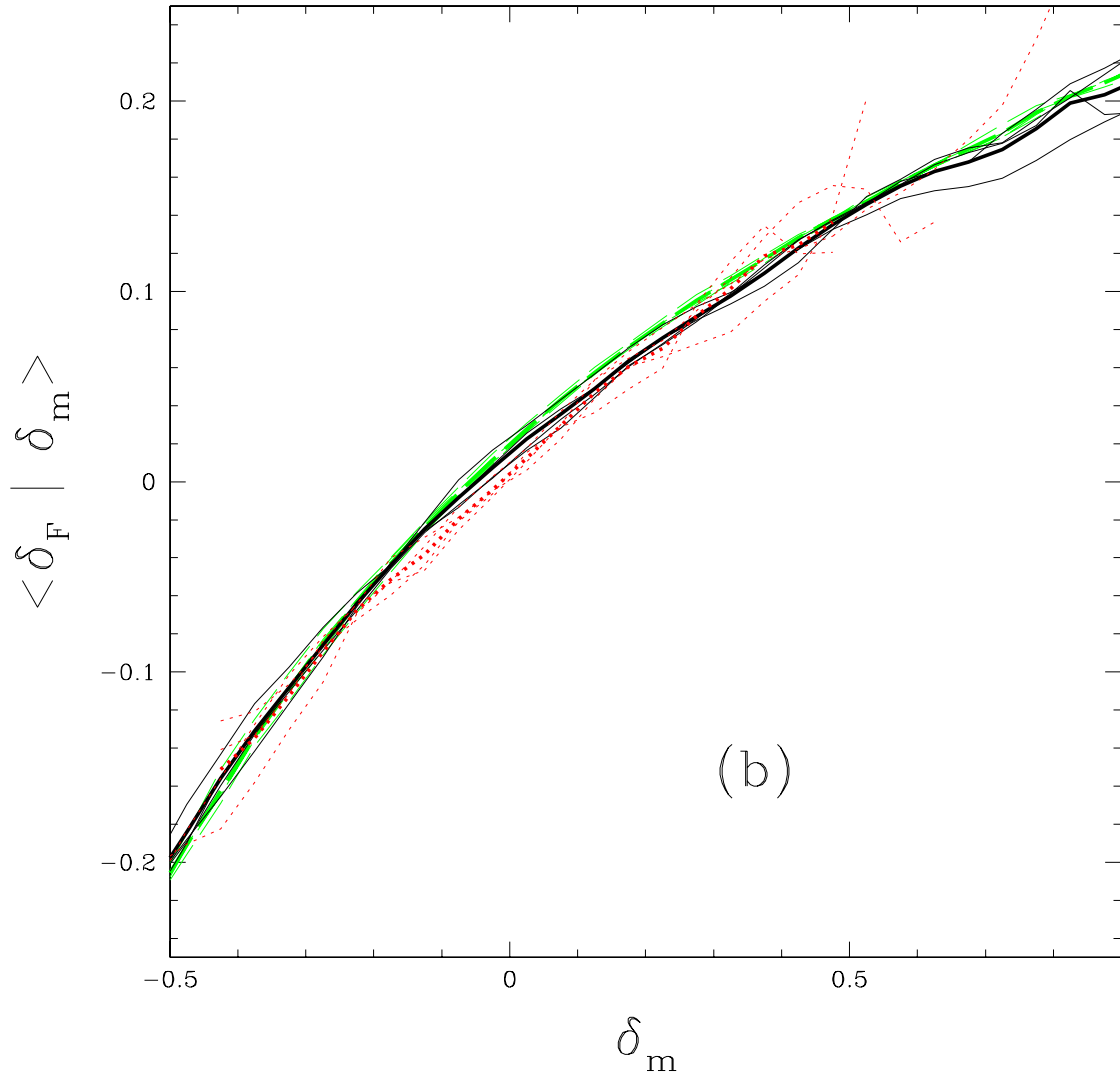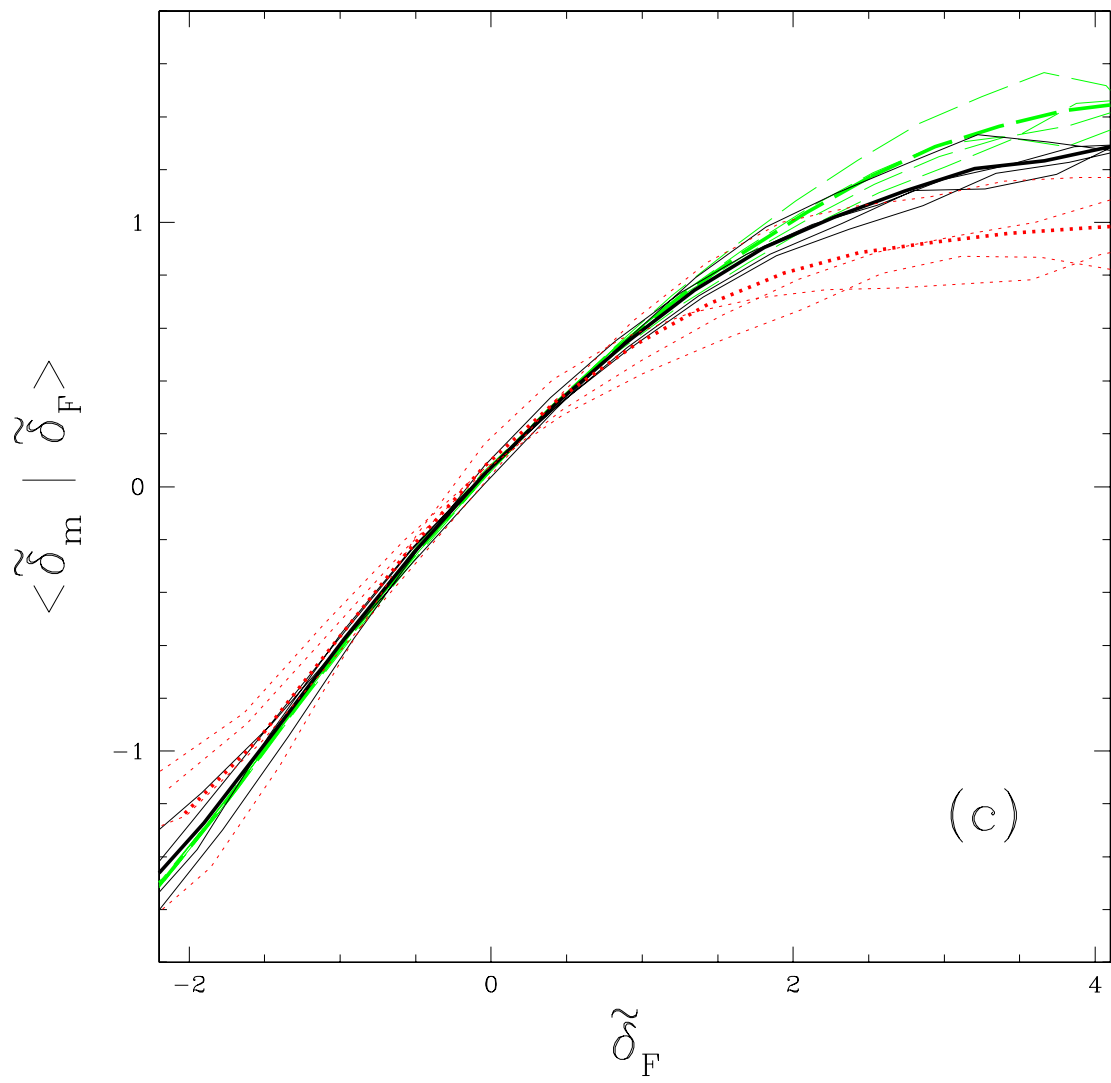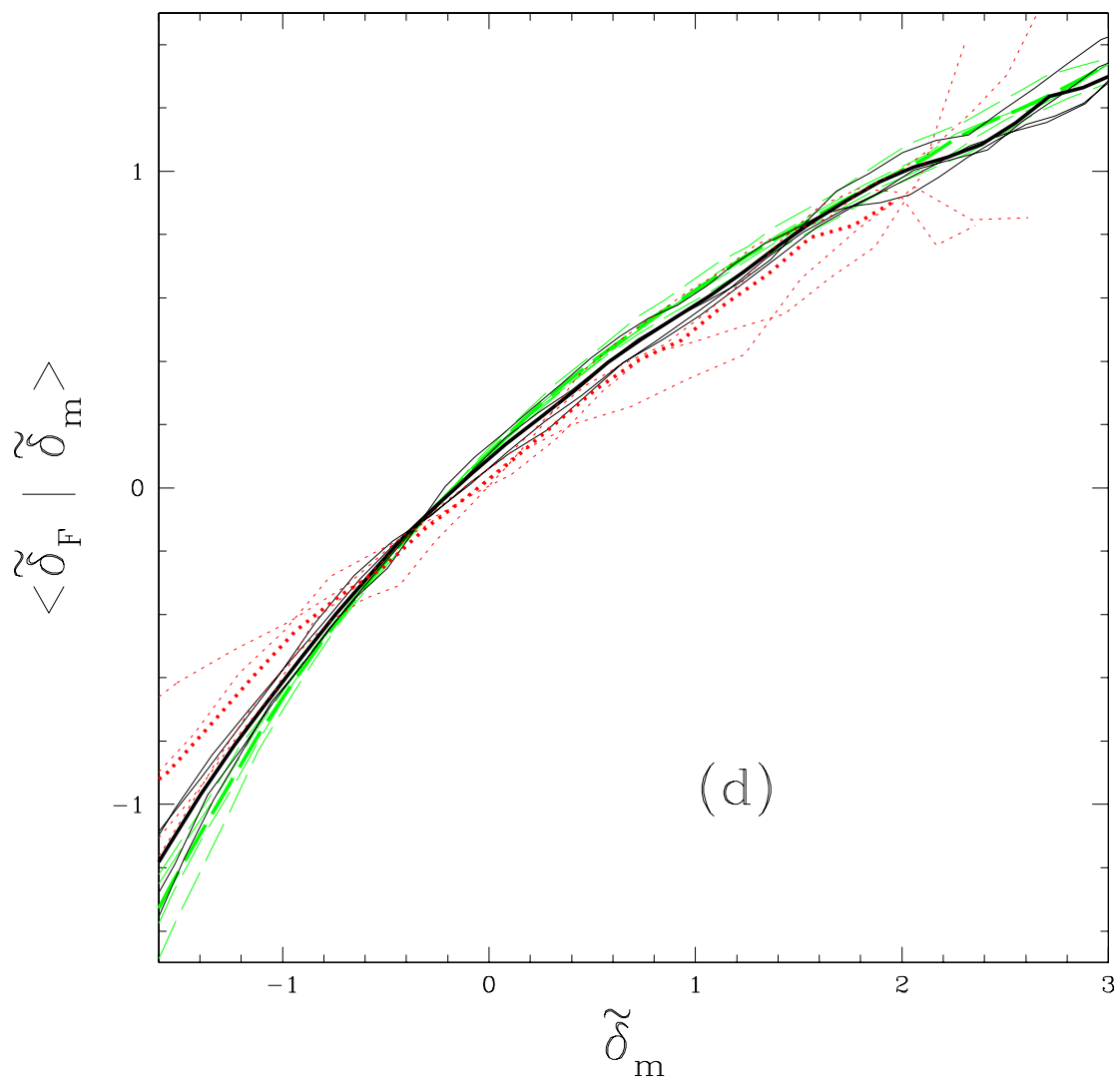
Fig. 5.— Box size and statistical error test. The thin curves are from four independent simulations, the thick curves are the average of the four. *Black (solid) lines:* $40\,h^{-1}$ Mpc simulations; *red (dotted) lines:* $20\,h^{-1}$ Mpc simulations; *green (dashed) lines:* $80\,h^{-1}$ Mpc simulations. All with $512^3$ particles (the $80\,h^{-1}$ Mpc simulations are slightly under-resolved).

(b)

fraction of the gas and thermal broadening.

## 3.1.  Mean transmitted flux

The value of the mean transmitted flux used in our standard model, $\bar{F} = 0.67$, is appropriate for $z = 3$. The variation of the mean transmitted flux is the most important factor accounting for the changes in the Ly$\alpha$ forest with redshift.

Figures 6(a-d) show the four functions defined previously for $\bar{F} = 0.67$ (black, solid line), $\bar{F} = 0.8$, appropriate for $z \simeq 2.5$ (green, long-dashed line), $\bar{F} = 0.6$, appropriate for $z \simeq 3.5$ (blue, short-dashed line), and $\bar{F} = 0.7$ (red, dotted line).

The functions $<\delta_m|\delta_F>$ and $<\delta_F|\delta_m>$ depend strongly on $\bar{F}$, whereas the functions with the normalized quantities $<\tilde{\delta}_m|\tilde{\delta}_F>$ and $<\tilde{\delta}_F|\tilde{\delta}_m>$ have a much weaker dependence. Essentially, the mean transmitted flux changes the effective "bias" in the Ly$\alpha$ forest that determines the value of $\delta_F$ that corresponds to a given mass fluctuation, and once a linear bias is eliminated the dependence on $\bar{F}$ is much smaller. Note that Figures 6a,b reflect the fact that the Ly$\alpha$ forest bias is larger for higher mean flux decrements (see Figure 10c of McDonald 2002, and Croft et al. 1999, McDonald et al. 2000).

We give the results for $\bar{F} = 0.6$ and $\bar{F} = 0.8$ in Table 1, as rows 6sd and 6ld, respectively. The first row in the Table already contains the results for $\bar{F} = 0.67$.

## 3.2.  Temperature-density relation

The temperature-density relation has practically no effect on the functions we are analyzing, as shown in Figures 7(a-d).     The reason for this insensitivity is that the temperature-density relation does not affect the large-scale properties of the Ly$\alpha$ forest directly, and apparently its effect through the large-scale bias factor is quite small (see Figure 10b in McDonald 2002).

## 3.3.  Amplitude of the power spectrum

In Figure 8, we show the variations with the amplitude of the mass power spectrum by increasing the amplitude in our standard model, shown by the solid line, by 33%, to produce the dotted line. In practice, we do this by simply using an output of the same simulation at a scale factor 1.33 times larger, but keeping the mean transmitted flux in the Ly$\alpha$ spectrum constant (this is completely equivalent to running a new simulation with a different amplitude, except that the effective temperature changes due to the change in the relation between comoving distance and velocity, but the previous figures show that the temperature does not affect the results).     The
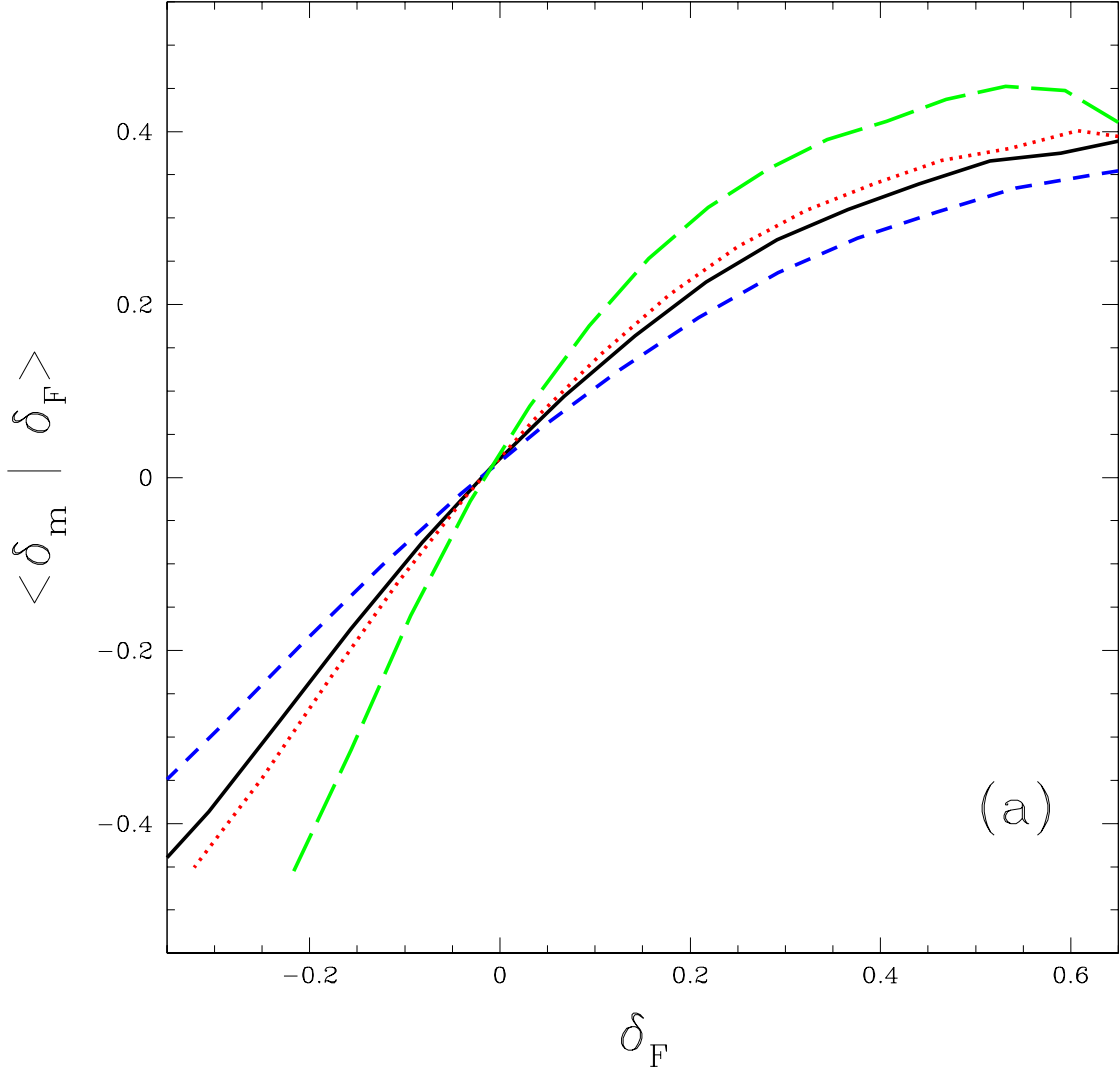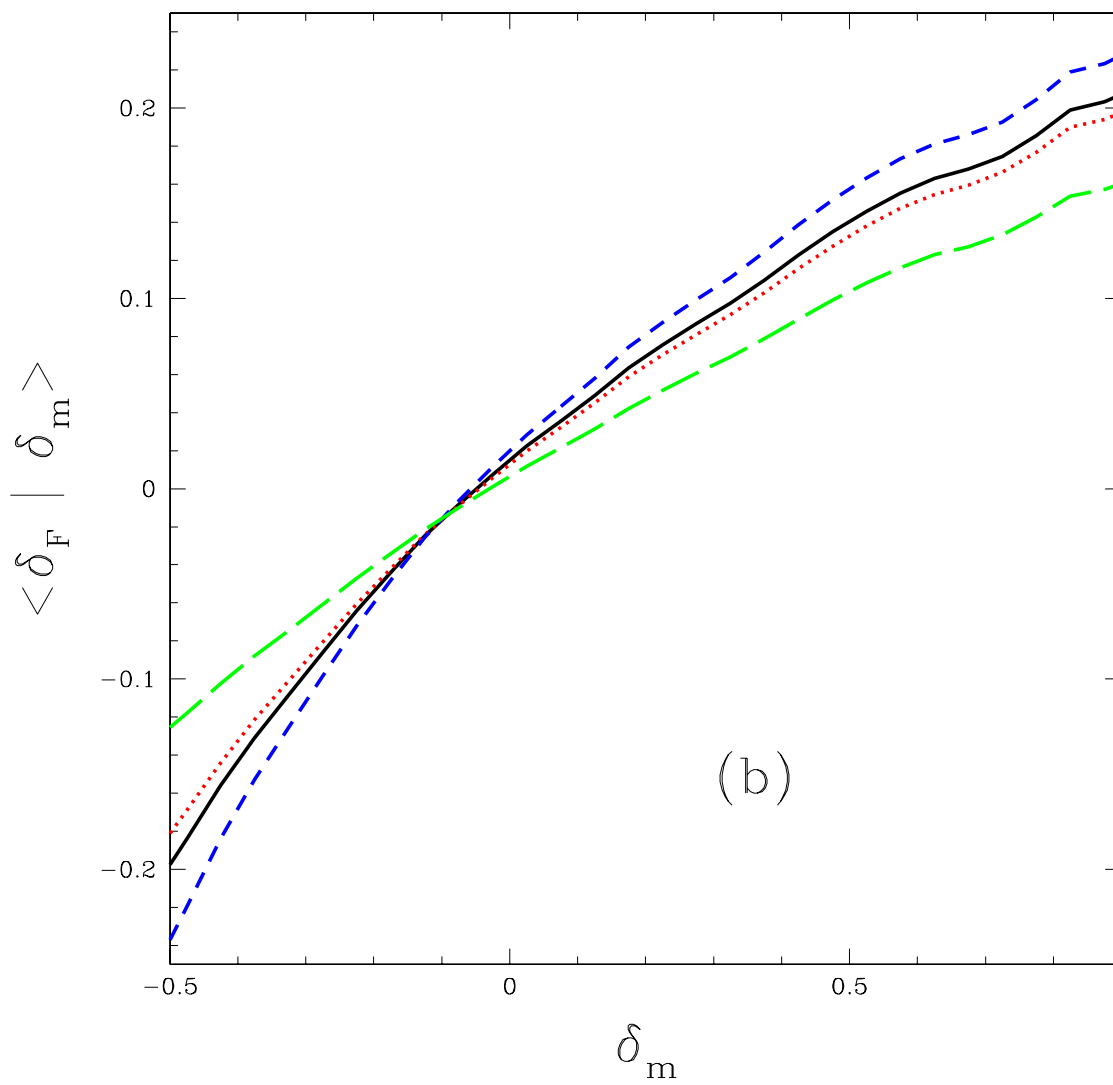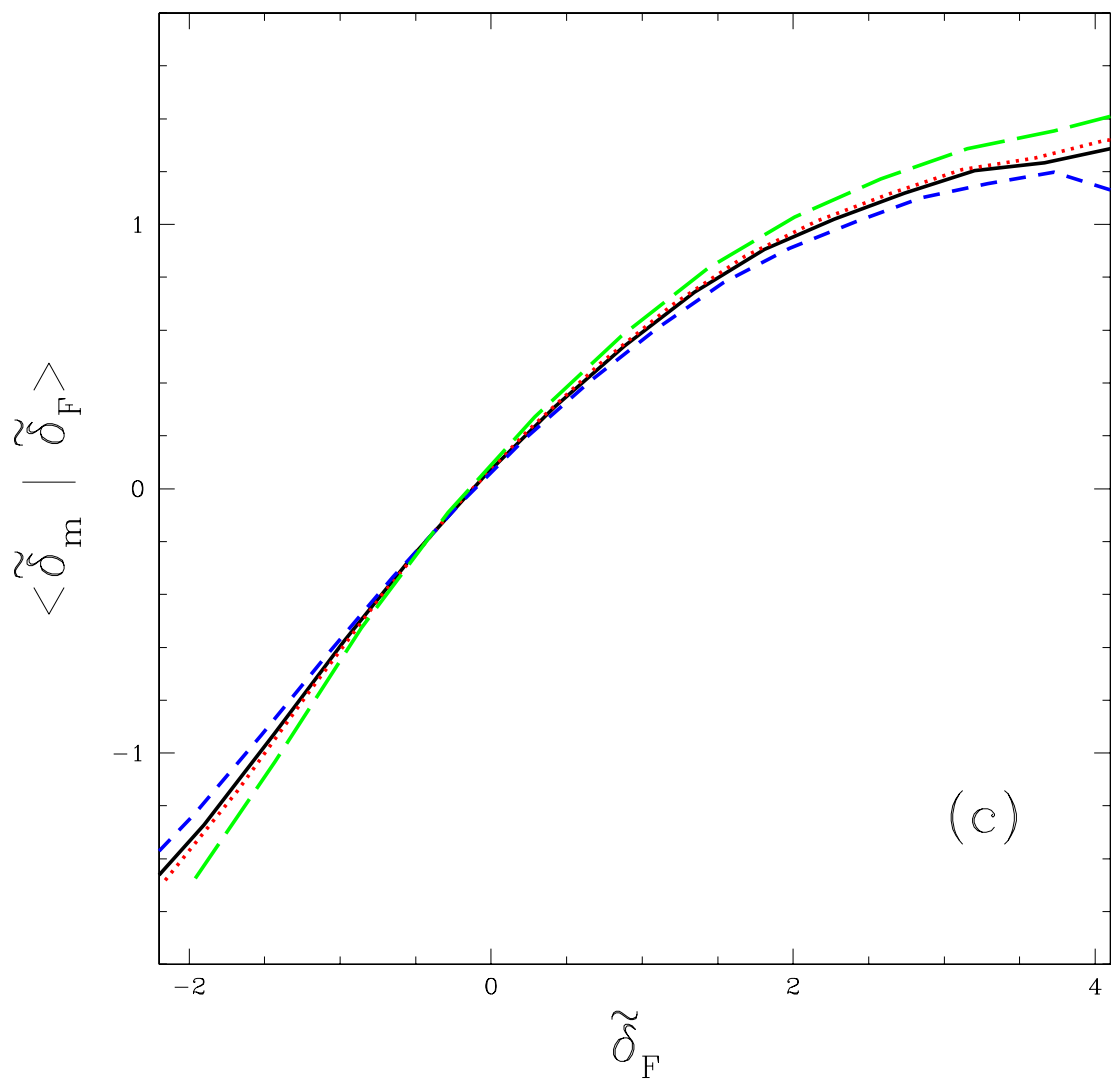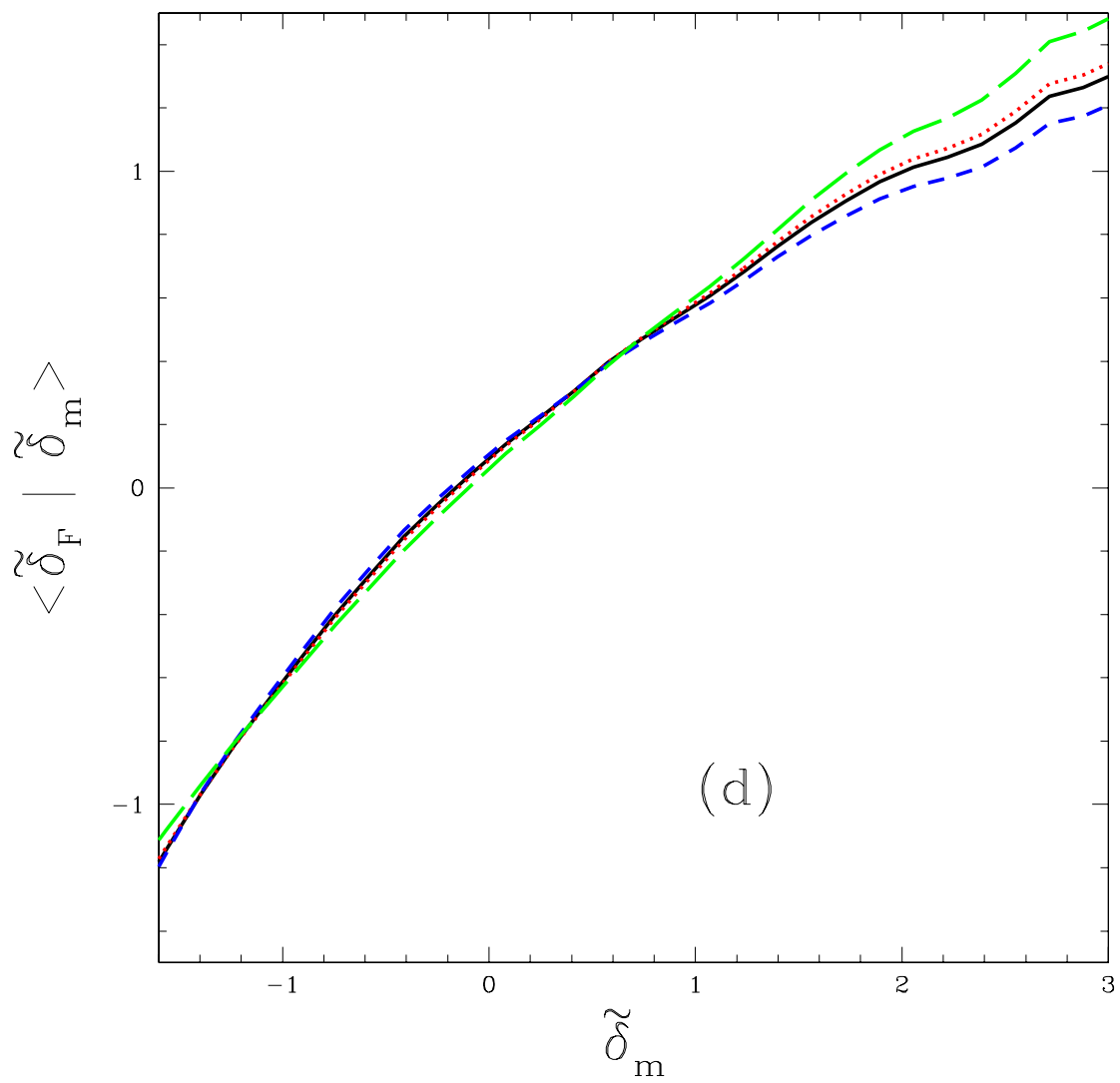
Fig. 6.— Effect of mean transmitted flux fraction. Shown are $\bar{F} = 0.67$ (black, solid line), $\bar{F} = 0.8$, (green, long-dashed line), $\bar{F} = 0.6$, (blue, short-dashed line), and $\bar{F} = 0.7$ (red, dotted line), for the same four functions as previously.
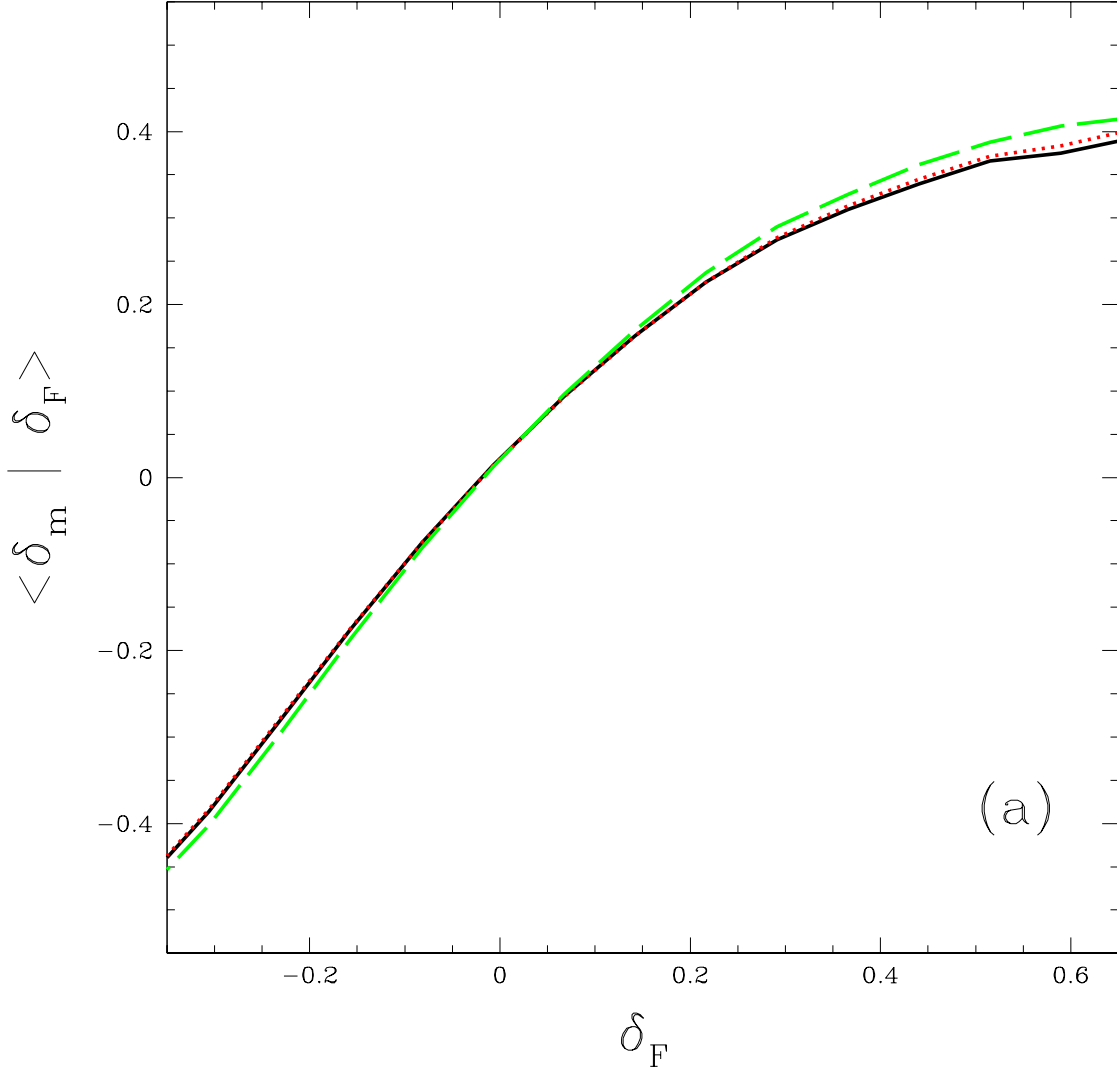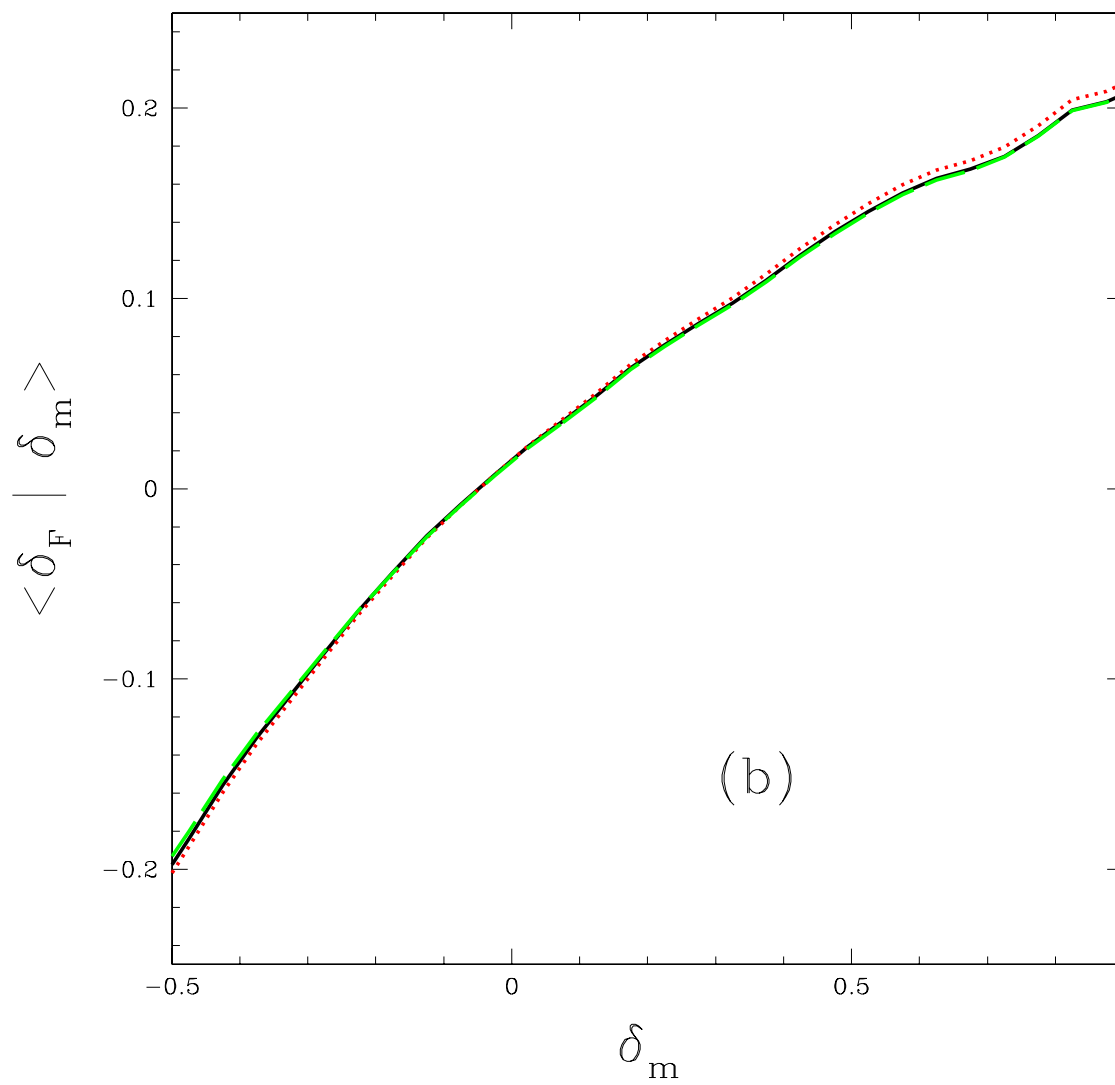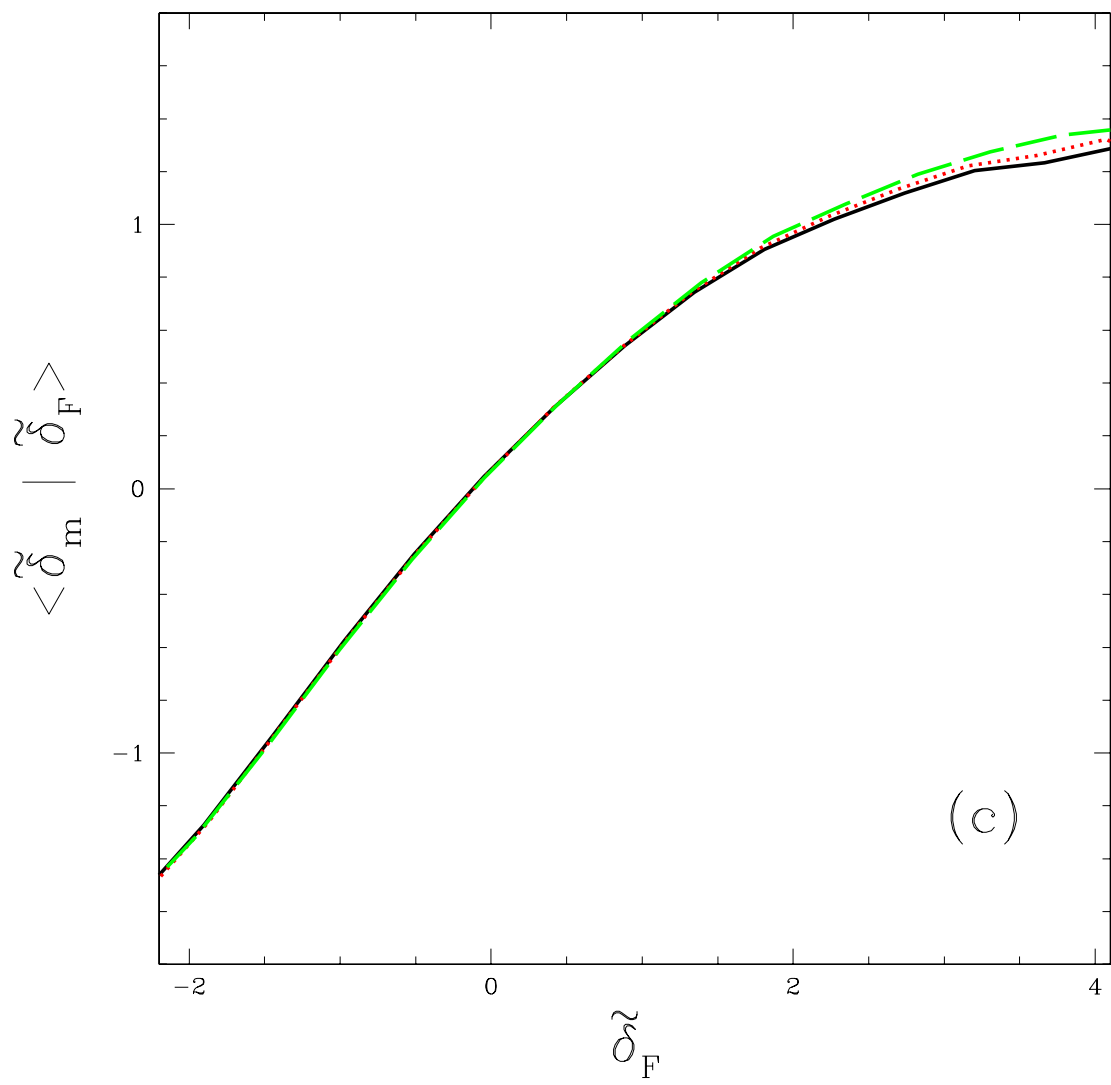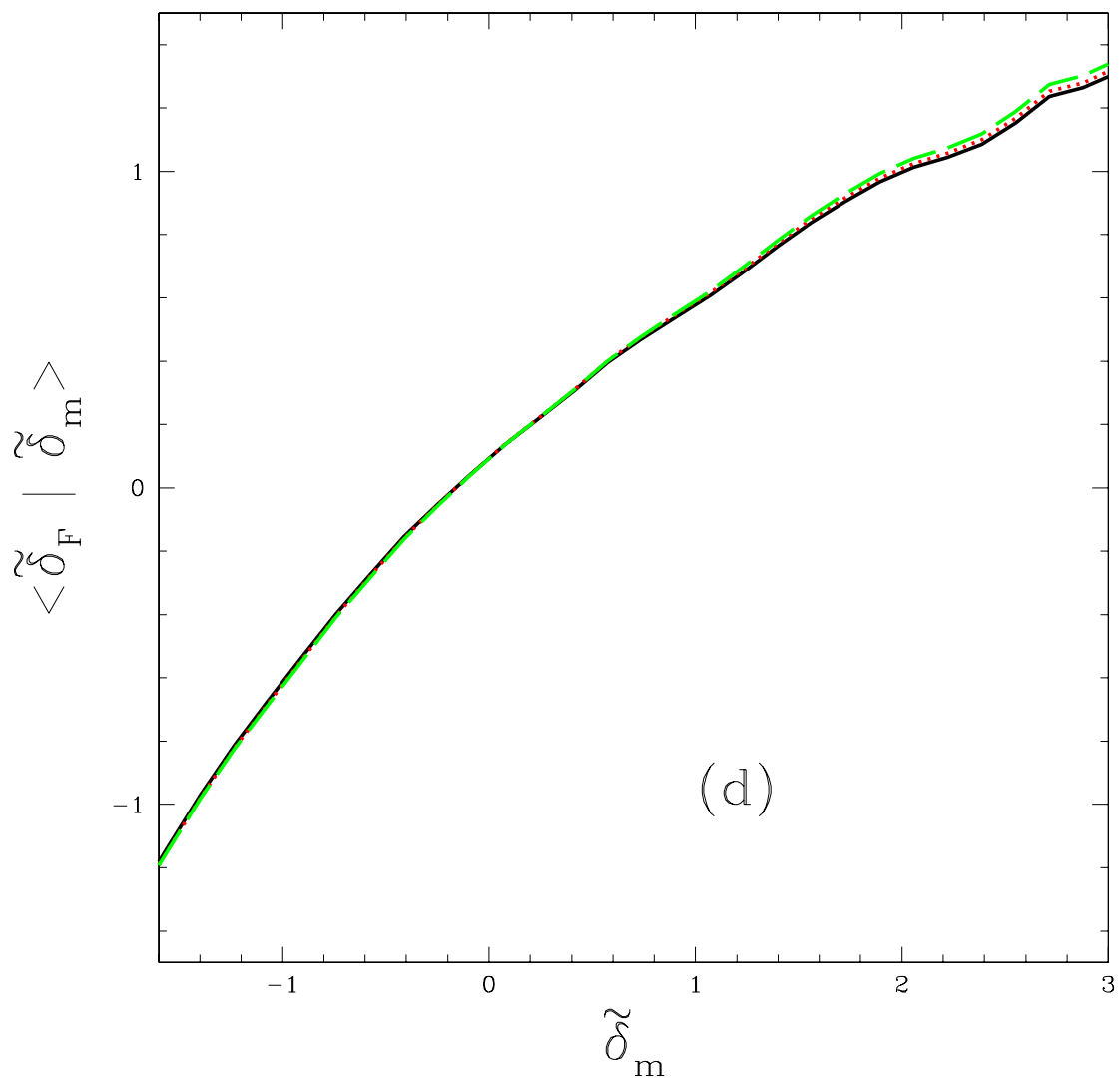
(b)

Fig. 7.— Effect of temperature-density relation. *Black, solid line:* $T_{1.4} = 17000$ K, $\gamma - 1 = 0.3$ (standard values); *red, dotted line:* $T_{1.4} = 22000$ K, $\gamma - 1 = 0.3$; *green, dashed line:* $T_{1.4} = 17000$ K, $\gamma - 1 = 0.6$.
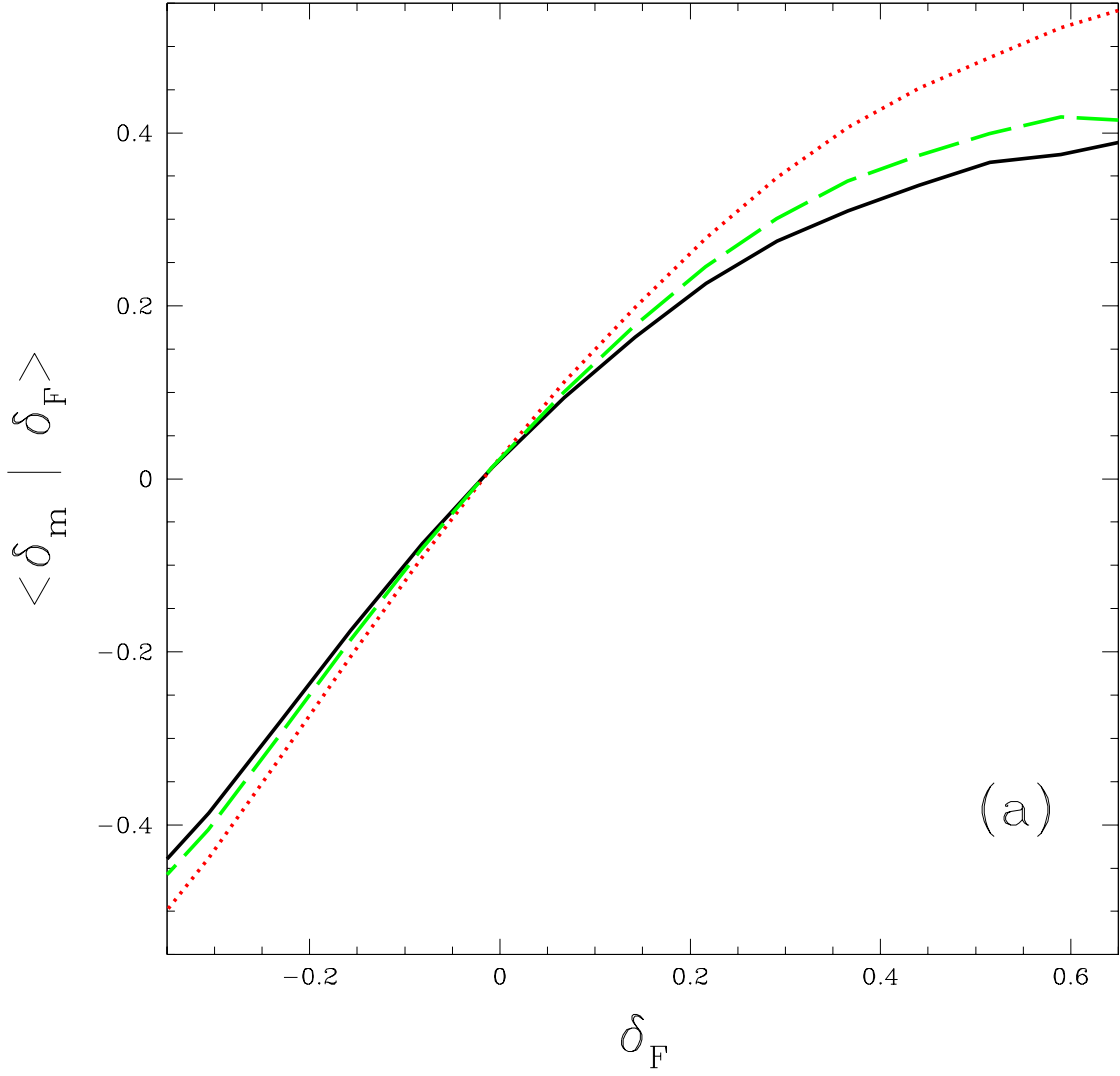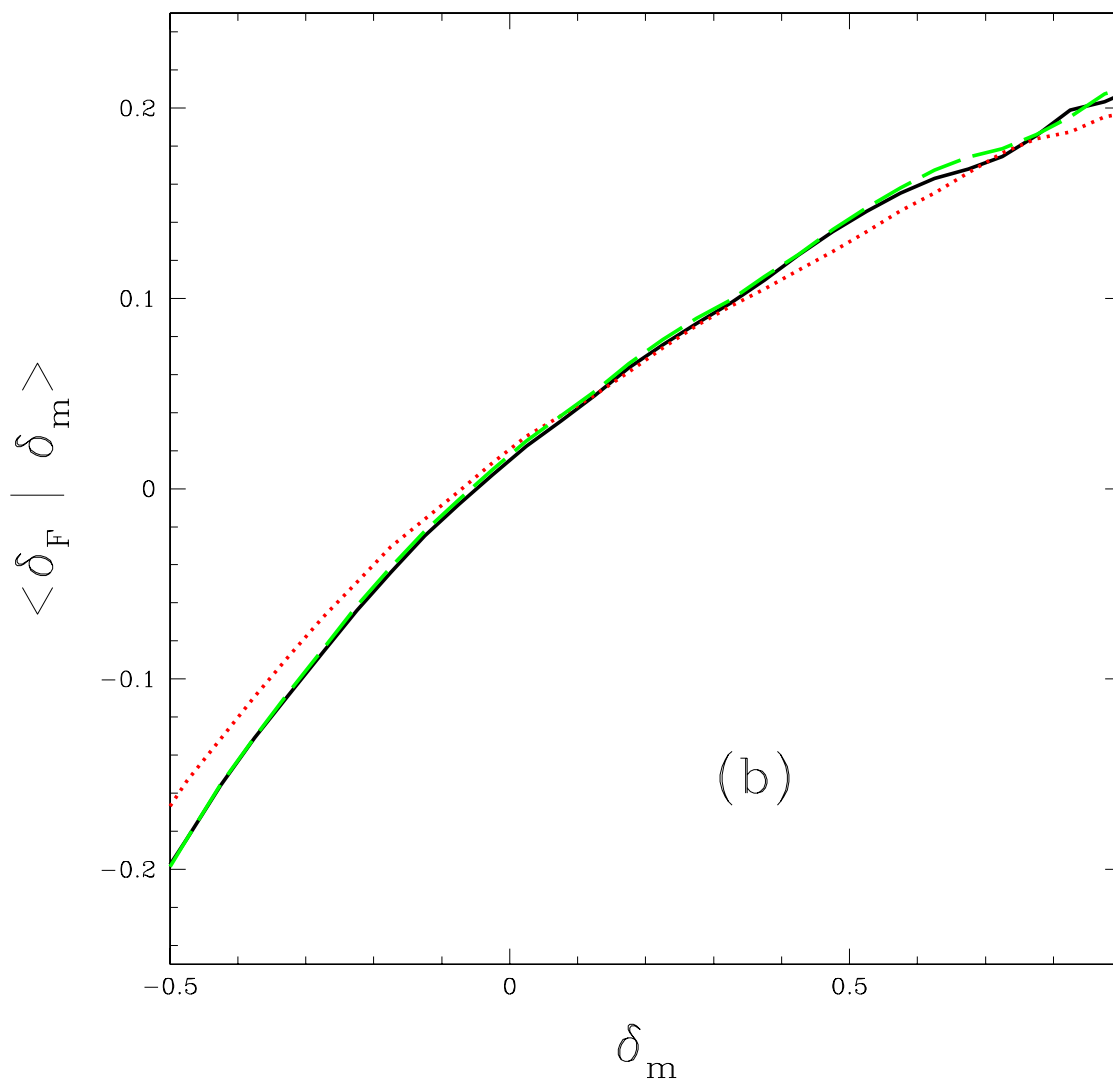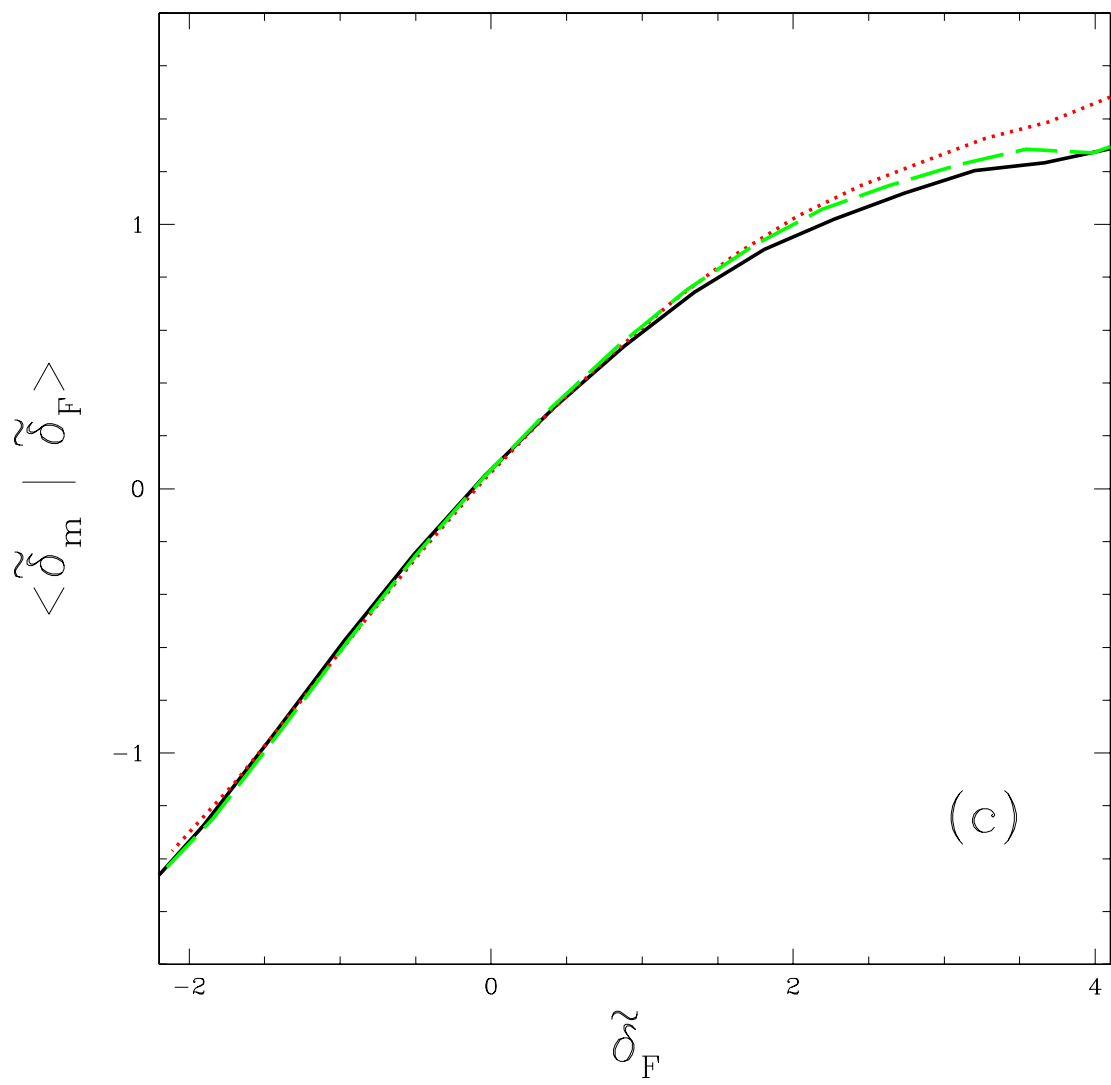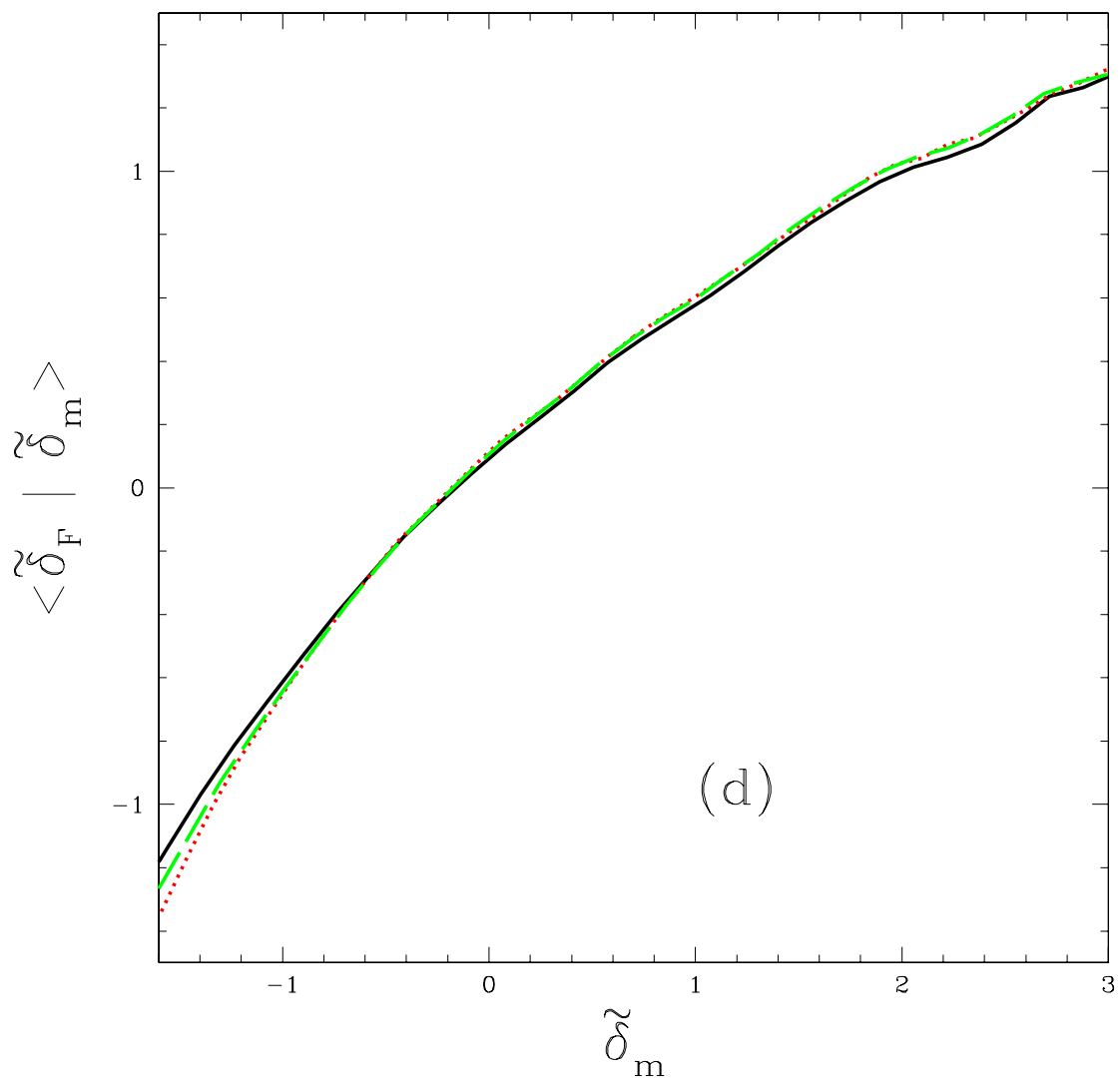
(b)

Fig. 8.— Effect of the mass power spectrum. *Black, solid line:* Standard power spectrum. *Red, dotted line:* rms amplitude increased by 33%. *Green, dashed line:* $n = 0.85$ instead of $n = 0.95$, with fixed amplitude at $k = 1 \ (h^{-1}\,\mathrm{Mpc})^{-1}$.

(d)

result reflects that the bias factor of the Ly$\alpha$ forest decreases with increasing power spectrum amplitude (see Figure 10a of McDonald 2002). The difference is partially cancelled when using the normalized variables $\tilde{\delta}_F$ and $\tilde{\delta}_m$, although less so for the function $< \tilde{\delta}_F | \tilde{\delta}_m >$.

We also vary the power-law slope of the power spectrum in Figure 8, by changing $n$ from 0.95 to 0.85 while holding the power fixed at wavenumber $k = 1 \, h^{-1} \, \mathrm{Mpc}$. The effect we see is relatively small. Had we carefully chosen the value of $k$ at which we hold the power fixed to minimize the variation with $n$, the effect might be even smaller.

We give the results for $\sigma_8 = 1$ and $n = 0.85$ in Table 1, as rows 8d and 8ld, respectively. The first row in the Table is for the standard power spectrum.

## 4. DISCUSSION

We have presented predictions for the expected correlation of the mass and the Ly$\alpha$ forest transmitted flux, smoothed over a cube size of $\sim 10 \, h^{-1} \, \mathrm{Mpc}$. Our most basic result is that the relation between $< \tilde{\delta}_m | \tilde{\delta}_F >$ and $\tilde{\delta}_F$ is linear over most of the range of $\tilde{\delta}_F$ (excluding rare, high $\tilde{\delta}_F$ values), with a slope of 0.6 on this smoothing scale. We have shown that this correlation is not sensitive to the temperature-density relation of the gas. There is a dependence on the mean transmitted flux and the amplitude of the mass power spectrum, but these quantities are already measured to reasonable accuracy (Croft et al. 1999, McDonald et al. 2000). We have also shown that the predictions do not suffer from uncertainties due to the resolution of the simulations. Calculations with even larger boxes than used here will be desirable, because the uncertainties in the predictions due to the variance in the simulations and the suppression of the large-scale power may still be significant; however, this uncertainty is mostly isolated in the most rare, high density regions.

A similar relation holds between $< \tilde{\delta}_F | \tilde{\delta}_m >$ and $\tilde{\delta}_m$. We notice, however, that the observational determination of this other function will be affected by galaxy shot noise. The number of galaxies observed in a certain cube in redshift space containing a fixed mass is subject to shot noise, and this inevitably introduces a smoothing of the function $< \tilde{\delta}_F | \tilde{\delta}_m >$, which must be taken into account before any comparisons to our theoretical results are made (in contrast, galaxy shot noise does not change the average $< \delta_F | \delta_m >$, but it does alter the rms dispersion $\sigma_F$ needed to compute $\tilde{\delta}_F$). Measuring the effects of galaxy shot noise can also teach us useful information about how galaxies form, because galaxy shot noise does not generally need to be strictly described by Poisson statistics. For example, for a fixed mass contained within a cube, once a galaxy is found in the cube the probability to find others may be lower because some mass has already been used up by that galaxy.

Comparing these theoretical predictions with observations allows for several tests of the basic theory of the Ly$\alpha$ forest, and can reveal new information on the spatial distribution of the galaxies. The functions $< \tilde{\delta}_m | \tilde{\delta}_F >$ and $< \tilde{\delta}_F | \tilde{\delta}_m >$, which should not be affected by any linear galaxy bias,

provide a powerful test of the assumption that the basic framework assumed here for the nature of the Ly$\alpha$ forest is correct, and that galaxies trace the mass on large scales apart from linear bias. If the slope of $< \tilde{\delta}_m | \tilde{\delta}_F >$ as a function of $\tilde{\delta}_F$ were found to be larger than predicted, this would imply that the Ly$\alpha$ forest is much more closely associated with galaxies than is expected from their common correlation with the mass distribution. If the observed slope were smaller than predicted, it would indicate that the Ly$\alpha$ absorbing gas and the galaxies tend to avoid each other for some reason.

If the correlation of $\tilde{\delta}_m$ and $\tilde{\delta}_F$ is as expected, this will imply a strong confirmation of the basic model we have for the Ly$\alpha$ forest, and will justify using the Ly$\alpha$ forest as a predictor of the mass fluctuations. Comparing the predicted function $< \delta_m | \delta_F >$ with the observed $< \delta_g | \delta_F >$ will then yield the bias factor of any type of galaxies for which these observations can be made.

The correlation of galaxies and the Ly$\alpha$ forest can be measured as a function of scale. Our predicted value of 0.6 for the slope of $< \tilde{\delta}_m | \tilde{\delta}_F >$ as a function of $\tilde{\delta}_F$ should decrease with the smoothing scale, in a way that reflects the shape of the mass autocorrelation function. This will allow a precise test of the idea that the large-scale distributions of different types of galaxies differ only in a constant linear bias relative to the mass fluctuations.

# REFERENCES

Adelberger, et al. 2001, in preparation

Bechtold, J., Crotts, A. P. S., Duncan, R. C., & Fang, Y. 1994, ApJ, 437, L83

Bergeron, J., & Boissé, P. 1991, A& A, 243, 344

Bi, H. G. 1993, ApJ, 405, 479

Bi, H. G. & Davidsen, A. F. 1997, ApJ, 479, 523

Cen, R. Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, ApJ, 437, L9

Cen, R., Ostriker, J.P., Prochaska, J.X., Wolfe, A.M. 2001, in preparation

Chen, H.-W., Lanzetta, K. M., Webb, J. K., & Barcons, X. 2001, ApJ, 559, 654

Croft, R. A. C., Weinberg, D. H., Bolte, M., Burles, S., Hernquist, L., Katz, N., Kirkman, D., & Tytler, D. 2002, ApJ, submitted (astro-ph/0012324)

Croft, R. A. C., Weinberg, D. H., Pettini, M., Hernquist, L., & Katz, N. 1999, ApJ, 520, 1

Dinshaw, N., Impey, C. D., Foltz, C. B., Weymann, R. J., & Chaffee, F. H. 1994, ApJ, 437, L87

Dinshaw, N., Weymann, R. J., Impey, C. D., Foltz, C. B., Morris, S. L., & Ake, T. 1997 ApJ, 491, 45

Dolan, J. F., Michalitsianos, A. G., Nguyen, Q. T., & Hill, R. J. 2000, ApJ, 539, 111

Gnedin, N. Y. & Hui, L. 1998, MNRAS, 296, 44

Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J. 1996, ApJ, 457, L51

Lanzetta, K. M., Bowen, D. V., Tytler, D., & Webb, J. K. 1995, ApJ, 442, 538

López, S., Hagen, H.-J., & Reimers, D. 2000, A&A, 357, 37

McDonald, P. 2002, ApJ, submitted (astro-ph/0108064)

McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., Cen, R., & Ostriker, J. P. 2000, ApJ, 543, 1

McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A.,& Cen, R. 2002, ApJ, in press (astro-ph/0005553)

Miralda-Escudé, J., Cen, R., Ostriker, J. P., & Rauch, M. 1996, ApJ, 471, 582

Monier, E. M., Turnshek, D. A., & Hazard, C. 1999, ApJ, 522, 627

Penton, S. V., Stocke J. T., & Shull, J. M. 2001, ApJ, in press (astro-ph/0109277)

Petitjean, P., Surdej, J., Smette, A., Shaver, P., Mueket, J., & Remy, M. 1998, A&A, 334, L45

Primack, J. R. 2000, preprint (astro-ph/0007187)

Rauch, M., Miralda-Escudé, J., Sargent, W. L. W., Barlow, T. A., Weinberg, D. H., Hernquist, L., Katz, N., Cen, R., & Ostriker, J. P. 1997, ApJ, 489, 7

Steidel, C. C., Dickinson, M., & Persson, E. 1994, ApJ, 437, L75

Zhang, Y., Anninos, P., & Norman, M. L. 1995, ApJ, 453, L57

Zhang, Y., Meiksin, A., Anninos, P., & Norman, M., L. 1998, ApJ, 495, 63

Table 1.   Selected Results

| Fig., line | $\sigma_F$ | $\sigma_m$ | $< \tilde{\delta}_m \vert \tilde{\delta}_F >$ $\tilde{\delta}_F = -1.5$ | 0 | 1.5 | 3 | $< \tilde{\delta}_F \vert \tilde{\delta}_m >$ $\tilde{\delta}_m = -1$ | 0 | 1 | 2 |
|---|---|---|---|---|---|---|---|---|---|---|
| 2, s [a] | 0.16 | 0.30 | -0.97 | 0.07 | 0.80 | 1.17 | -0.61 | 0.09 | 0.58 | 1.00 |
| 2, ld [b] | 0.15 | 0.26 | -0.95 | 0.09 | 0.73 | 1.03 | -0.56 | 0.09 | 0.53 | 0.94 |
| 2, sd [c] | 0.18 | 0.36 | -1.00 | 0.05 | 0.87 | 1.35 | -0.68 | 0.12 | 0.63 | 1.05 |
| 6, ld [d] | 0.11 | 0.30 | -1.09 | 0.10 | 0.86 | 1.26 | -0.63 | 0.06 | 0.60 | 1.11 |
| 6, sd [e] | 0.19 | 0.30 | -0.92 | 0.06 | 0.77 | 1.12 | -0.60 | 0.11 | 0.56 | 0.94 |
| 8, d [f] | 0.18 | 0.39 | -0.98 | 0.06 | 0.84 | 1.27 | -0.65 | 0.12 | 0.60 | 1.03 |
| 8, ld [g] | 0.17 | 0.33 | -1.00 | 0.07 | 0.84 | 1.21 | -0.64 | 0.11 | 0.60 | 1.03 |

Note. — First column gives the Figure and curve to which the row of results applies. Solid line=s, dotted=d, long-dashed=ld, and short-dashed=sd. Value for $< \tilde{\delta}_m \vert \tilde{\delta}_F = 3 >$ underestimated due to limited box size. $\sigma_m$ and $\sigma_F$ intended only for use as conversion factors from $\tilde{\delta}$ to $\delta$ – alone they may be sensitive to resolution and box size.

[a]Standard case.

[b]$12\,h^{-1}$ Mpc cube, $4.2\,h^{-1}$ Mpc Gaussian.

[c]$8\,h^{-1}$ Mpc cube, $2.7\,h^{-1}$ Mpc Gaussian.

[d]$\bar{F} = 0.8$.

[e]$\bar{F} = 0.6$.

[f]$\sigma_8 = 1$

[g]$n = 0.85$.